Categorizing recipients of evouchers using best practises from marketing theories

Clustering and targeting vulnerable recipients of evouchers using a novel approach of consumer segmentation and machine learning; a case study of Sint Maarten

Delft University of Technology

D.G.M. (Daan) Gorsse

This page was intentionally left blank

Categorizing recipients of evouchers using best practises from marketing theories

Clustering and targeting vulnerable recipients of evouchers using a novel approach of consumer segmentation and machine learning; a case study of Sint Maarten

by

D.G.M. (Daan) Gorsse

Institution:	Delft University of Technology
Place:	Faculty of Technology, Policy and Management
Project Duration:	December, 2021 - May, 2022
Student Number:	4299736
Course:	EPA2942

To be defended publicly on Thursday June 2, 2022.

Graduation Committee

Chair & second supervisor:	Prof. Dr. M.E. (Martijn) Warnier	Multi-Actor Systems
First Supervisor:	Prof. Dr. T.C. (Tina) Comes	Engineering Systems and Services
Advisor:	Dr. Y. (Ylenia) Casali	Engineering Systems and Services
External supervisor:	Ir. T.G.J. (Tijs) Ziere	510, Netherlands Red Cross
External supervisor:	Dr. M.J.C. (Marc) van den Homberg	510, Netherlands Red Cross

An electronic version of this thesis is available at http://repository.tudelft.nl/.
Associated code and models are available at: https://github.com/DGMGo/MasterThesis.git.

Cover Image: Recipients of an evoucher project conducted by the Red Cross (Netherlands Red Cross, 2021c)

Delft University of Technology

An initiative of the Netherlands Red Cross

D.G.M. Gorsse: *Categorizing recipients of evouchers using best practises from marketing theories* (2022)

Preface

Dear reader,

During my years in Delft, I discovered many different topics which I could explore during my master thesis. Humanitarian aid was the one that sparked my interest the most, since I cannot think of a better way to help more people in need in a short period of time. Last half year I have had the opportunity to dive deeper into the field of Cash and Voucher Assistance (CVA). CVA is growing rapidly and with it, the need to further organize related information. Unfortunately, it is a fact that due to climate change (natural) disasters will increase in frequency and intensity. CVA will contribute to mitigating some of the negative consequences while ensuring dignity of aid recipients. Since the humanitarian sector has big aspirations with CVA in the future, it is my hope to be able to help making these ambitions a reality.

As with most master theses, this junior researcher went through several formative moments during this research process. At times it was definitely a challenge to conduct data science research without having a sound background in data science. Luckily, this challenge was paired with a lot of fun because of the novelty of the research domain and the potential impact this research can have in the humanitarian sector. The humanitarian sector lags behind in process innovations compared to the faster and more advanced commercial sector. One of the personal learning points of this research is that I found that it works best for me to visualize a final product on the horizon and works towards this end goal. Explorative research is another cup of tea since you should not start with the best method, but you just must chose one and build something while tinkering on.

This research was not possible without the help of my dear graduation committee. Tina, you were the first person from whom I heard about humanitarian aid during my studies. This inspired me to become more involved in this very fascinating sector. This shows that passionate teachers can trigger the 'next generation' to be involved helping less fortunate people. I am also grateful for your razor-sharp view and constructive criticism during the early formation of my research. Ylenia, thanks a lot for the numerous meetings where I could discuss many of the dozens of impossible research methods I came up with. Your scientific view, theoretical insights and thorough feedback brought this thesis to a higher level. Martijn, although we did not see each other frequently, I benefited greatly from your broad experience and calming advise. Thanks for this and for your reassuring positive outlook! Marc, I am grateful that this journey at 510 started with an early LinkedIn reply from you with the recommendation to maybe do something with CVA. Without your connections to other people in the field and your suggestions, this research would have been less extensive. Finishing with the one and only Tijs, I want to show my appreciation for your contagious enthusiasm, realism and support. It would be nice if we could work more on improving CVA in the future! Also thanks to people who facilitated this research with sometimes smaller or bigger additions to the puzzle. Martina from the Erasmus University had relevant input on her expertise on marketing theories. And for the people from 510 and the Netherlands Red Cross who I could use as sparring partners: Lars, Jonath, Jacopo and Sabrina. I am looking forward into working with them even more in this great organisation with a worthy mission!

Last, but certainly not least, thanks to my close friends and family who had to leave me studying for quite some evenings and weekends. Especially my girlfriend Lisanne who is one of the few who could watch the process from up close. Thank you for your patience and conversations where we countlessly restructured this research. My elaborate time in Delft ends with this research. I am grateful for getting the opportunity to develop myself on so many levels during my time in Delft. My hope for you is that you enjoy reading this research and that you feel inspired by it.

D.G.M. (Daan) Gorsse
The Hague, May 2022

Abstract

Cash and Voucher Assistance (CVA), a type of humanitarian aid consisting of giving money instead of products, is being used more frequently because of its effectiveness and efficiency in helping people in need (Cash Learning Partnership, 2020b). The debate on using CVA is currently focusing on improving the quality by better incorporating 'voices' (needs and preferences) of recipients and by enhancing targeting. In targeting it is a major challenge to quickly identify the individuals and families with the biggest needs, given the lack of data (Aiken et al., 2021).

Research on ways of measuring impact on and satisfaction of recipients combined with research on demographic and behavioural characteristics of recipients could lead to deeper insights in recipients of trackable CVA modalities (evouchers and ecash). This research uses the marketing literature on customer segmentation combined with machine learning algorithms to come up with an innovative new approach of categorizing recipients of evouchers, using the case of a Red Cross project on Sint Maarten. The main research question is: How can recipients of cash and voucher assistance be categorized using the field of consumer segmentation by using machine learning methods?

The objective of this research is to come up with new methods to better understand recipients of CVA. Theories on customer segmentation pointed to the use of data-driven clustering methods to categorize consumers. Combined with a framework of recency, frequency and monetary aspects, recipients of evouchers could be categorized effectively. A required addition to this clustering method is to use a dimension reduction technique to avoid the negative consequence of the curse of dimensionality. Therefore, a two-step approach of dimension reduction and clustering has been applied in this research.

It has been found in this research that a factor-cluster approach can lead to insightful clusters using geo-demographic data and behaviour data. Factor analysis has been used to reduce the dimensions while the k-prototype algorithm has been used to cluster into five distinct groups of recipients. The geo-demographic variables that were the most determining in characterizing distinct clusters consisted of: the age of the main beneficiary, the different household compositions and of a constructed factor 'big families and big receivers'. The most distinguishable variables on behaviour were: the number of supermarket visits (frequency), the time between the first voucher was received and the first transaction (recency) and the variables on the amount of money that was spent with the vouchers (monetary).

To be able include the 'voice' of recipients (needs and preferences), a connection between the registration data, behavioural data and survey data is needed. In this research only an exploratory connection could be established, due to the lack of a common identifier between the survey data and the other datasets. However, one crucial finding of this research is that it seems like the combination of these data sources can give meaningful insights in the needs, preferences and behaviour of households of Sint Maarten. With these insights specific clusters can be targeted for additional assistance, based on their needs.

Recommendations for future studies include studying the validity of the found cluster results with different validation indices on cohesion and compactness, and by using simulations to determine the cluster stability. Before this factor-cluster approach can be deployed in CVA projects, more research on the treatment of limitations of this approach needs to be conducted. This is critical in communicating the conditions and constraints of this model to humanitarian aid workers in the field. Another recommendation is to improve the design of surveys to measure the needs of recipients. For insightful factor-cluster results on the needs of recipients, survey data should be linked to geo-demographic and behaviour data. More research on including clusters in retargeting methods using feedback loops have a large potential in minimizing targeting errors and more effectively meeting the needs of recipients. With this research, the humanitarian sector can benefit from new ways to understand the needs of the most vulnerable in need. Decision-makers should build upon the feedback of recipients and move towards a new era of humanitarian assistance.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
CaLP	Cash Learning Partnership
CARE	Cooperative for Assistance and Relief Everywhere
CVA	Cash and Voucher Assistance
CTP	Cash Transfer Programming
DG ECHO	Directorate-General for European Civil Protection and Humanitarian Aid Operations
EPA	Engineering and Policy Analysis
FSP	Financial Service Provider
IARAN	Inter-Agency Research and Analysis Network
ICRC	International Committee of the Red Cross
IFRC	The International Federation of Red Cross and Red Crescent Societies
MEAL	Monitoring, Evaluation, Accountability and Learning
MPC	Multipurpose Cash
NAF	Netherlands Antilles Guilder
NGO	Non-Governmental organisation
NLRC	Netherlands Red Cross
NRRP	National Recovery and Resilience Plan
ODI	Overseas Development Institute
PDM	Post distribution monitoring
RCRCM	Red Cross and Red Crescent Movement
SXM	Sint Maarten
UN OCHA	The United Nations Office for the Coordination of Humanitarian Affairs
VfM	Value for Money

List of Figures

List of Tables

1

Introduction

This chapter aims to describe the problem and knowledge gaps that are the context of this research project. The objective, main research question and sub questions are presented in section 1.2 followed by a description of the research approach.

1.1. Problem definition

Defining the problem gives the relevant background information of the issue at hand. Diving into Cash and Voucher Assistance reveals the current and future complexity of problems related to including the needs of users.

1.1.1. Increasing need of humanitarian aid

With the rise of climate change worldwide, we can expect more and more natural disasters to happen in the coming decades. Disasters such as floods, forest fires and hurricanes will become more common in parts of the world because of more extreme weather patterns and rising sea levels (IPCC, 2021). Also, seemingly slower natural disasters, such as extreme drought, will lead to more emergency situations. Most of the casualties of natural disasters are happening in poorer countries (Norton et al., 2020). Especially, vulnerable communities in these countries will be more affected and often (extreme) poverty is the determining factor here (Ritchie & Roser, 2020). Humanitarian aid is often needed to help and support these affected communities. The United Nations Office for the Coordination of Humanitarian Affairs (OCHA) expects an increase in the need for humanitarian aid in the coming decades (2021). In 2021, an estimated 1 in 33 people worldwide needed help, compared to 1 in 45 people in 2020, which already was one of the highest ratios in the last decades. Different kinds of aid can be given by a large variety of organisations. Providing health care, temporary shelter, clean water and food packages were and still are a very important way to support people in need.

1.1.2. The promise of Cash and Voucher Assistance

One particular type of humanitarian aid increasingly used is Cash and Voucher Assistance (CVA). CVA is a different type of aid in the way that it comprises giving cash or vouchers (documented and restricted money) instead of distributing products to people. This type of humanitarian aid is a means to a specific objective and not a goal in itself. Objectives differ from sector-specific goals such as providing food during emergencies to multipurpose cash transfers for a variety of possible expenditures. A key difference of CVA compared to in-kind products is that is allows people to act as economic agents ensuring 'dignity of choice' (Vogel et al., 2022). During the World Humanitarian Summit in 2016, international donors, aid organisations and governments of multiple countries have pledged in the so-called Grand Bargain to "increase the use and coordination of cash-based programming" (World Humanitarian Summit, 2016). One of the specific agreements that was made is that at least 20 percent of the budget of aid should be given to local organisations to make humanitarian interventions as local as possible. Local capacity is crucial in improving the effectiveness of aid projects since local organisations have the best understanding of needs and contexts (Nightingale, 2012). Other commitments from the 2016 Grand Bargain are to invest in new delivery models of aid and to build an evidence base to assess the costs, benefits, impacts and risks of CVA. Furthermore, far-reaching collaboration and coordination regarding CVA should be achieved to develop standards, guidelines and monitoring and evaluation mechanisms. After the Grand Bargain, the United Nations Inter-Agency Standing Committee published the 'Strategic note on Cash Transfers in Humanitarian Contexts' in cooperation with the World Bank Group (2016). The European Civil Protection and Humanitarian Aid Operations (ECHO) followed with their 'Guidance to deliver large-scale cash transfers' in 2017 (ECHO, 2017). These ambitions, among many other efforts, lead to an increase in cash used for CVA programmes worldwide from an estimated 2.8 billion US dollars in 2016 to 4.5 billion US dollars in 2018 and 5.6 billion US dollars in 2019 (Cash Learning Partnership, 2020b). Together with these growing budgets, the scope of projects is also increasing. Coming from small pilots with hundreds of recipients of CVA in 2004, the largest project in the world is reaching 1.5 million refugees in the Emergency Social Safety Net in Turkey (Cuevas et al., 2019).

1.1.3. Delivering better quality CVA

Although the use of CVA is growing, it has not yet reached its full potential according to The Cash Learning Partnership (2020b). Currently, around 18 percent of total global humanitarian aid expenditures is reserved for CVA projects (Development Initiatives, 2021b). There are many organisations with big aspirations and targets on making CVA more standardised and giving it a more prominent role in humanitarian aid. It is promising that major aid organisations such as the International Red Cross and Red Crescent Movement made it one of their flagship projects to scale up their cash and voucher assistance to at least 50 percent of humanitarian assistance by 2025 (IFRC, 2020). The phrase "Why not cash?" has been used more and more often since the High Level Panel on Humanitarian Cash Transfers placed more emphasis on starting with considering CVA instead of other modalities (High Level Panel on Humanitarian Cash Transfers, 2015). Bailey in 2016 is one of the first researchers who started her research by asking "Why not cash?" when considering aid modalities for refugees in Mozambique (2016). With these influences, the debate of the last 10 years has moved from "Why cash?" to "Why not cash?" and currently in the direction of "How to scale up cash and how to do better quality CVA?". To achieve better quality CVA, opinions and satisfaction users of CVA should be included more and there should be more ways to measure the impact of CVA. This frequently heard ambition will be further clarified and divided in the next two sections.

1.1.4. Shortcomings of excluding users' voices

One of the aspects falling behind with the advance of CVA is the design and evaluation from a user-centred perspective. Ground Truth Solutions systematically analysed feedback from over 10,000 crisis-affected individuals from all over the world (2018). They find that recipients of CVA generally do not feel included. In the report of an outlook of CVA in the future to 2030, CaLP and IARAN (Inter-Agency Research and Analysis Network) also state that recipients of CVA do not feel listened to and that there are no clear signs of improvement for the future (2019). So, they argue that the needs and voices of users should be placed at the centre and "humanitarian actors need to urgently explore to increase accountability to users". They further stress the need for aid organisations to "transparently capture what works and scaling only the most effective models, while ensuring a strong user voice in this process". Burton, currently working as ICRC's Institutional Lead for CVA, further elaborates on the need to actively listen and reflect on their preferences and needs, ensuring a person-centred approach to humanitarian aid (IFRC, 2020).

Ground Truth Solutions expresses their concern for the most vulnerable in need, if donors and practitioners of CVA are not able to shift their mindset from "more CVA" to "better CVA". To maximise the promise of choice, dignity and value of cash programmes, recipients' needs and preferences should be consulted (Ground Truth Solutions, 2019). In order to do better quality cash and voucher assistance, we should move towards a more user-centred approach when doing research on, implementing and evaluating CVA projects. Listening to and taking into account the needs and preferences of recipients is crucial in bringing CVA to a higher quality where opportunities and dignity of vulnerable people can be restored, as set as objectives of CVA on the World Humanitarian Summit (2016).

Many existing evaluations of CVA projects are done from a process- and cost-perspective instead of a user-perspective. From conversations with The Red Cross Netherlands, one of the possible reasons for this, is the influence of donors who want to see numbers on costs and results. Organisations in the humanitarian aid sector are focused on reaching higher targets regarding CVA projects and on being visible to get resources from donors (Peachey, 2021). Metrics on how much money is spent can tell something about efficiency of a project but does not reveal much on effectiveness and user satisfaction. It is harder to monitor the impact on recipients and how their needs and preferences are considered. Obrecht points to the concept of path-dependency when explaining the reserved willingness and hesitancy of donors to invest in CVA (2017). Path-dependency is the concept that most decision-making processes are following a more or less fixed trajectory based on previous decisions with a small range of different future diversions. Humanitarian decision-making in the literature has seen more examples of path dependency slowing decision-making (Darcy et al., 2013). This makes it a challenge for donors to accept the relatively new CVA projects and this can act as a “barrier to evidence- and data-driven resource allocation” (Obrecht, 2017). CaLP states that although there are many comparative cost-efficiency and cost-effectiveness studies, the evidence base of CVA “remains weak and fragmented” (Cash Learning Partnership, 2020b). In conclusion, more research on including the users’ voices is needed. This can result in new ways of measuring the effectiveness of CVA projects and improving the quality of CVA projects.

1.1.5. Targeting accuracy

When implementing CVA projects, targeting is one of the policies where users’ voices can be implemented and can have a positive influence on the effectiveness and efficiency of the humanitarian aid. Targeting comprises the study on the needs, preferences and behaviour of users to find new ways to better prioritise aid. One of the most crucial aspects of a successful and efficient food security or social safety net intervention is accurate targeting (Barrett et al., 2013). Current targeting methods are expensive, constrained by resources and time-consuming (Verme & Gigliarano, 2019). Also, the accuracy of targeting varies greatly in humanitarian aid projects with reviews stating that in various projects around 25% of the targeting performed worse than random allocation of resources (Altındağ et al., 2021). Hanna and Olken argue that targeting errors are a major source of inefficiency in humanitarian aid (2018). Accurate targeting can be achieved by minimizing leakages (reaching people who do not need aid) and by minimizing undercoverage (not reaching people who are in need). Accurate targeting is frequently hindered by a lack of institutional capacity and credible data. As a result, available aid is distributed through a variety of proxy mechanisms, including simple ways like demographic or geographic targeting, as well as more advanced mechanisms like community or self targeting (Altındağ et al., 2021).

The fundamental challenge with targeting people in need is to rapidly identify households and individuals with the biggest needs, often with limited data. Recent data of populations is hardly available, for example 50 percent of the poorest countries performed a census in the last 10 years (Yeh et al., 2020). Humanitarian workers should be informed by the needs of targeted groups, but this is currently lacking in many projects due to the lack of relevant data. Targeting can serve two purposes: nowcasting and contemporaneous prediction (Browne et al., 2021). Nowcasting combines recent observations of predictive variables with previous observations to support in targeting and geographic needs analysis, and also providing baseline metrics for impact evaluations. Contemporaneous prediction is used to fill in gaps not covered by the available data, for example with absence of locations in surveys to predict poverty estimates of a specific geographic area (Browne et al., 2021). A promising spot on the horizon is to evolve targeting to real-time monitoring of CVA projects to quickly fill in the gaps in needs of people in need (CaLP & IARAN, 2019; Moreno, 2017).

Recommendations for improvement from researchers point in the direction of coupling targeting with behavioural data. For example, details on consumption may be useful to support targeting. There are existing aid projects conducting expenditure and consumption surveys, but they are time-consuming (McBride & Nichols, 2016). New sources of data on behaviour can help to accurately target in future projects, for example by the use of segmentation analysis. By identifying the distinct features of customer groups, we can better understand them and deliver tailored aid to each segment (Daellenbach et al., 2018). Segmentation can help in identifying segments who may be more vulnerable than others and therefore require more or different assistance. Socioeconomic features, when combined with an analysis of people’s purchasing behaviour, may provide organizations with useful information for targeting vulnerable groups and obtaining information that might possibly amplify the beneficial impact of a cash transfer (Van den Homberg et al., 2019). The Cash Learning Partnership stresses the importance of improving targeting approaches by new methods of data analytics and ensuring a strong user voice in the process (CaLP & IARAN, 2019).

1.1.6. Research needed to do more accurate targeting including users

There are specific knowledge gaps in the literature that can be possibly addressed using this case from a user-centred perspective in order to improve targeting during CVA projects.

Measuring the impact on and the satisfaction of recipients

CaLP notes in their State of the World's Cash Report of 2018 that "there is a perception that the quality of CTP [Cash Transfer Programming] is improving, but there are no objective measurements of quality" (CaLP, 2018). It is important to know how well humanitarian aid organisations listen to recipients and to be able to measure this based on real data. Deciding for the recipients on what they need or want to consume will result in a mismatch between the help provided and the real demand based on early research (Gelan, 2006). To guarantee that the goods or services given meet the genuine needs of the recipients, a more evidence-based approach is required in humanitarian aid (Castillo, 2021). There are some key performance indicators on CVA related to processes, but not to users needs, preferences and satisfaction. Key performance indicators based on efficiency and other money related outcomes exist and are often used in cash projects (IRC, 2016). Accountability refers to how well organisations listen to the needs of recipients and other stakeholders. For example, recipients of aid can point to room for improvement, or they can complain and make their needs more clear during aid projects. According to the Red Cross Netherlands, this is currently one of the most difficult activities to execute and measure this in the sector. Measuring accountability is often overlooked and sometimes even entangled by other seemingly synonyms like effectiveness. The Directorate-General for European Civil Protection and Humanitarian Aid Operations states that accountability considerations should use recipient satisfactions, coping strategies and different perceptions of well-being (2015). But it is yet unknown how to incorporate satisfaction, needs and preference when monitoring and evaluating a CVA project. To better measure and monitor the quality of CVA projects, key performance indicators related to recipients are needed, such as including beneficiaries' satisfaction rate, needs, preferences and other associated measurements. In conclusion, there are existing key performance indicators to measure and monitor effectiveness in CVA projects, but it is yet unknown how to measure beneficiaries' needs, preferences and satisfaction. Also, the relation between satisfaction of recipients and impact of the program is not yet researched and could be correlated.

Influence of household compositions and recipient characteristics

There are certain characteristics and compositions of households which are not yet fully understood in CVA projects. Particularly vulnerable people such as people with disabilities, single-headed households, elderly, refugees and child-headed households are still a very important overlooked group according to the Grand Bargain 2.0 endorsed framework (The Inter-Agency Standing Committee., 2021). Especially in Sint Maarten supporting the vulnerable is important since their economy is still recovering while essential reforms in public financial management, tax reforms and improved social safety nets are urgently needed (International Monetary Fund, 2021). Also, in the evaluation of the CVA project on Sint Maarten, it was pointed out that accessibility to people with disabilities could be improved and that the "registration system was not the most adequate to ensure that all vulnerable people had easy access to registration" (Menéndez & Barrena, 2021).

According to CARE International and ideas42 (2020), women and female-headed families are typically disadvantaged since their environments are often more restrictive, meaning that financial assets and coping strategies available to them are limited. As a result, their positions in aid projects are more fragile and vulnerable. Another interesting group to dive into are close communities or subgroups of recipients. There are assumptions that there is different spending behaviour and perspectives on risk by different communities in the case of Sint Maarten. Also, Ojiambo and Chamaa consider that various communities react differently to the risks of delivery mechanisms of CVA (2021).

When assessing a new project demanding humanitarian aid, the modality Cash and Voucher Assistance is often preferred by the people affected compared to in kind or other services (CaLP, 2018). But this is not always the case since preferences can, for example, be influenced by the perceived value of support, market volatility, and expertise with a specific modality (CaLP, 2018). There is no research yet on the difference in preferences of recipients on what impact this made on their satisfaction and needs fulfilment. Before starting the CVA project on Sint Maarten, around 20 percent of the eligible people for CVA did not prefer CVA over in-kind assistance. In the evaluation of this project interviews were conducted to explore how recipients perceived evouchers above unconditional cash (Menéndez &

Barrena, 2021). The freedom of choice in the supermarkets by the evouchers was perceived positively while the opinions were mixed on the potential use of unconditional cash. Although a lot of literature on CVA states that unconditional cash is often the better modality, opinions from recipients are divided and dependent on context. For example, this preference can also flip to the extreme other side, which was the case during a CVA project in Gaza where only 30 percent of the beneficiaries preferred CVA as it was viewed as unreliable and too easy to discontinue in future phases of the aid project (Key Aid Consulting., 2018).

Recipient behaviour when using CVA

There exists a knowledge gap regarding behaviour of recipients and what impact this has on CVA projects. For example, there is assumed emergent behaviour in spending cash and vouchers. Hoarding, panic-buying and stockpiling could occur in a system where there is limited availability of items and a high demand combined with a fear of scarcity (Y. Chen et al., 2020). Saving, defined here as not being influenced by others, can also have downsides on the system during a CVA project. This was visible at certain times during the CVA project in Sint Maarten, as reported by the Red Cross. Other behaviour of recipients which is useful to dive into is selling of items obtained with CVA. There are examples of recipients selling food and non-food to obtain money for more pressing needs including health care, debt repayment, and education. This is inefficient since recipients seldom receive fair market value for their items, and sales from relief supply can destabilise local markets by promoting commodity hoarding and generating price fluctuations (MercyCorps, 2007). In a known case in Kenya at a refugee camp, around 12 percent of the refugee's income consisted from reselling food rations (IFC, 2018). Next to selling, the perception exists that CVA could have potential connection with fraud, corruption and misuse of CVA. Although this risk is debunked multiple times in the literature, the evaluation of the project on Sint Maarten mentions several people from interviews expressing their concern that these vouchers would be misused for purchasing alcohol, cigarettes or lottery tickets (Cash Learning Partnership, 2020b; Menéndez & Barrena, 2021). In conclusion, behaviour of recipients regarding stockpiling, panic-buying, hoarding, saving and selling is assumed to have a significant impact on recipients. The relationship with targeting recipients and their behaviour should be further researched in order to better understand behaviour of recipients of CVA.

1.1.7. Problem statement

Summarizing the previous sections, CVA is a promising and increasingly applied type of humanitarian aid. However, the quality of CVA needs to be improved, especially by better including recipients' voices in CVA projects. An aspect of CVA in which these voices can be incorporated are targeting approaches used in CVA projects. In the research field of CVA there is a need for more research on:

- Including the impact on and the satisfaction of recipients
- Including the influence of household compositions and recipient characteristics
- Including recipient behaviour

This research project covers the above aspects by exploring the use of consumer segmentation theory and machine learning models to improve future user-centered targeting in CVA projects. The positioning of this research is visualized in figure 1.1.

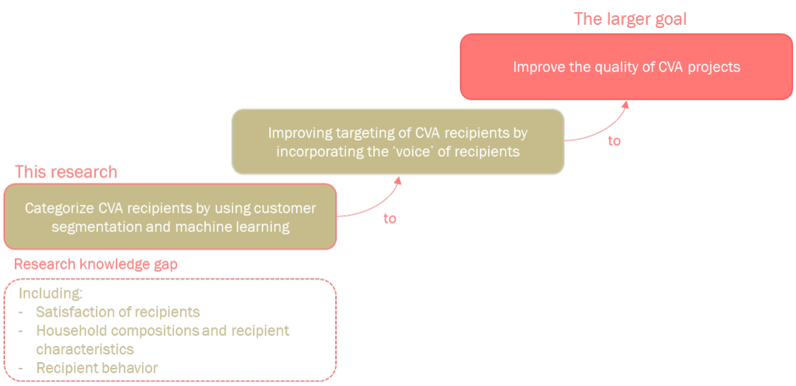


Figure 1.1: The research in context of improving the quality of CVA

1.2. Research setup

The research setup explains the objectives, main research question and sub questions.

1.2.1. Objective

The objective of this research is to come up with new methods to better understand recipients of cash and voucher assistance. This is crucial for the improvement of accurate targeting, which will be done by calibrating supply with factual and current needs for better tailored assistance. Giving voices to the voiceless will be crucial in increasing understanding the needs, preferences and behaviour of recipients of CVA projects.

The novelty of this exploratory research lies in the application of behaviour data, unsupervised machine learning methods (e.g., clustering) combined with existing research on marketing theory. Literature on segmentation of customers is widely available and has been researched extensively since the 1950's and can potentially help in handling customer behaviour data in research on CVA. The combination of behaviour data, customer segmentation theories and clustering has not been applied before in the humanitarian sector on CVA and could potentially open new doors to better targeting. By reaching this objective, the hope is that aid organisations can reach their ambitions by prioritising the poorest, the most in need and most vulnerable people and ensuring that they receive enough aid. This will contribute to doing better quality CVA in the future.

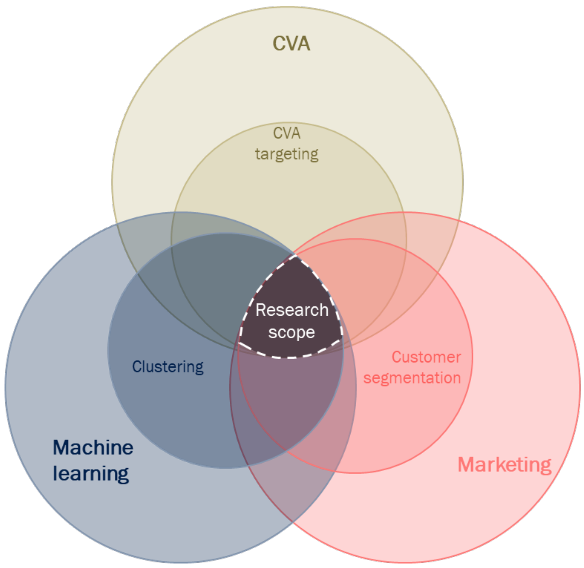


Figure 1.2: The research fields combined in this research

Proposed outcome

While this research will be applied to the case of Sint Maarten, the goal is to generalise this method to apply it on similar evoucher projects. Some outcomes could be derived from this research. First, one of the outcomes could be a new framework of how you can conduct real-time analysis of evouchers to better target vulnerable people in the upcoming phases of CVA projects. This could be done by identifying useful segments in the affected population with a clustering method supported by developed theories on marketing. Secondly, a possible outcome could be a list on which data might be useful to collect in a CVA project to analyse recipients needs using the earlier mentioned research on consumers and spending behaviour. Vice versa, a list of which data should not be collected could be important too because of protecting privacy rights by data minimization.

Thirdly, it could be an outcome to be able to state, for example, that a specific recipient has a certain probability of needing extra help based on previous spending behaviour and the current household composition. This extra need could be to assist this recipient by more frequent communication, by more or different allocation of evouchers, by personal consultation or by delivering (a combination of) different aid. There are numerous possibilities how aid organisations can help people with specific needs. The ideal future for humanitarian aid workers is to know in real-time which (groups of) people need extra help so aid workers can help them by a quick reallocation of the right resources (CVA or other aid) to a specific person or group.

1.2.2. Main research question

The use of digital cash and evouchers (and the amount of adjacent data) is growing and makes way for new possibilities for steering CVA projects by targeting and helping vulnerable people more effectively. Especially the different household compositions, characteristics, needs and behaviour are not being understood enough. There is a lot of existing research on (the combination of) spending behaviour with (big) data analytics, real-time monitoring and consumer characteristics. While this is mostly used in the commercial sector, for the humanitarian sector this is a novelty and can be used for improving projects on cash and voucher assistance (CVA).

The main research question that will lead this project is:

How can recipients of cash and voucher assistance be categorized using consumer segmentation and machine learning methods?

1.2.3. Sub questions

The following sub questions are presented to help answering the main research question:

Chapter 2 on Cash and Voucher Assistance

- What are common CVA configurations?
- How are CVA projects carried out?
- What are frequently used policies to target and help people in need?

Chapter 3 on Case Study: Sint Maarten

- What is the timeline of the CVA project on Sint Maarten and which actors are involved in the decision-making and setup of the project?
- What are the first interesting observations of this evoucher project and why is this a relevant case?
- What data is available for analysing recipients in our case study?

Chapter 4 on Data Preparation

- How can the datasets be cleaned and combined?
- What does exploratory data analysis tell us about this dataset?

Chapter 5 on Customer Segmentation

- What customer segmentation methods are potentially applicable to this research?
- What are variables from the field of customer segmentation that can be useful in describing behaviour of CVA recipients?
- How can these methods be applied on the data of Sint Maarten?

Chapter 6 on Modelling

- What are important aspects to consider when selecting segmentation models?
- What models are there available and do these compare?
- How can the model be evaluated and validated?

Chapter 7 on Results

- Which factors can be found using factor analysis?
- Which cluster characteristics are most distinguishable?
- What insights can survey data reveal when using the cluster results?

Chapter 8 on Conclusion

- What are the limitations of this study and how are they affecting the outcome?
- How can the outcomes of this study be generalised to other crises where CVA projects are used?
- How can recipients of cash and voucher assistance be categorized using the field of consumer segmentation by using machine learning methods?
- What are recommendations for humanitarian aid workers and academics?

1.3. Research approach

The approach of this research will be of a prescriptive and explanatory nature since the goal of this research is to give more insight to humanitarian aid workers to enable better quality CVA in the future. Exploratory research is used to address novel challenges for which there has been little or no prior investigation. Using methods of market basket analysis and consumer segmentation have not been done before in the humanitarian sector.

1.3.1. Case study research

The case study approach is beneficial to this research since one of the aims of this research is to explain abstract phenomena and a specific location provides context and data. That makes this research more realistic by coupling data to the model using real historic data and a specific location provides context and data which can be used for theory building (Eisenhardt, 1989). Conclusions should be interpreted in the context of the environment of Sint Maarten or similar regions, while they cannot be applied generally to any geographical area. When acknowledging the limitations and assumptions with this case study, there is the possibility to compare to other cases to generalise certain conclusions. The case of the CVA project on Sint Maarten is particularly useful since there is useful data available and there is a possibility to contact the relevant actors. This study will be carried out to assist the Netherlands Red Cross's 510 team.

The Red Cross Netherlands, part of The International Red Cross and Red Crescent Movement, has been conducting different CVA projects in the Country of Sint Maarten, often abbreviated to SXM. The island of Sint Maarten was heavily affected by the hurricane Irma in 2017 and is still recovering from the damage (Van der Mee, 2020). The Covid crisis put their citizens in a new phase of disaster since St. Maarten is heavily dependent on income from tourists. The real GDP of Sint Maarten in 2019 was lower than 2010 (International Monetary Fund, 2021). The Netherlands Red Cross started different aid projects, such as an evoucher project which ran from 2019 to 2021. They are looking for ways to do better quality CVA for possible future projects. For most humanitarian aid organisations, it is a challenge to target and prioritise the needs of the most vulnerable. The mission of the Red Cross is to help the most vulnerable first and the CVA project on Sint Maarten was a means to reach this goal (Netherlands Red Cross, 2018). This case will be presented extensively in chapter 3.

Using a specific case for evaluating CVA projects is a good way for researching on how to improve CVA. Especially people from the Caribbean are not being heard well enough according to CaLP (Cash Learning Parthership, 2020). Compared to other global clusters, there is limited evidence for CVA in the Caribbean and this has "restricted the ability of evidence-based decision-making on the allocation of limited resources". Furthermore, CaLP stresses the need to evaluate projects to ensure that CVA is still effective in meeting the various and changing needs of affected populations. Furthermore, analysis of past projects can be used for advocacy and to acquire new stakeholder buy-in (Cash Learning Parthership, 2020).

This CVA project on Sint Maarten is unique because there is an extensive amount of behaviour data available on vouchers and demographics of recipients. The Grand Bargain Workstream states that granular accessible data on transfer values is still scarce and could be an improvement if made publicly accessible (Development Initiatives., 2021a). There has been a project evaluation from a qualitative point of view but not yet from a quantitative view using this unique dataset. Furthermore, they presume that some interesting emergent recipient behaviour happened such as saving, stockpiling and that there are differences in spending patterns in subgroups and characteristics of recipients. Also, the people on Sint Maarten had to deal with different kinds of disasters, such as an abrupt shock by hurricanes and lingering stresses from the current Covid crisis. This makes Sint Maarten a very interesting case to use for researching CVA and to help decision-makers of CVA projects in the future with doing better quality CVA in different kinds of disasters.

1.3.2. Innovative methods

This research is unique in the application of two innovative methods, namely the use of consumer segmentation and by the use of machine learning methods.

Consumer segmentation

The field of consumer marketing can help with customer segmentation and market basket analysis to help to accurately target people. Marketers frequently utilize cluster analysis to create market segments, which allows for better product and messaging placement. Companies utilize this to better position themselves, explore new markets, and produce and sell goods that are relevant and valuable to certain clusters. The already developed literature on this could enhance the analysis of existing and future CVA projects. There are no examples of research using these fields combined with CVA, especially since there is limited amount of data on behaviour available.

Machine learning models

Machine learning (ML) models can help with quickly analysing big amounts of data and observing patterns. Classification and clustering are two ML methods that classify large structures into smaller segments; this helps to identify unexpected similarities between vulnerable people in need (Moreno, 2017). Other methods include making profiles of households whose needs are not being met sufficiently. There is a small number of papers using ML methods in the humanitarian sector (Aiken et al., 2021; Monaghan & Lycett, 2013). They showed that with machine learning models and new (big) data sources, patterns of poverty can be recognized, and new targeting approaches can be constructed and applied in various fields in the humanitarian sector. Furthermore, in humanitarian emergencies where time is scarce, ML algorithms can be a useful support to decision making (Knippenberg et al., 2017).

1.3.3. Demarcation

This research will be on evouchers in the form of value vouchers which is a specific modality of cash and voucher assistance. These evouchers on Sint Maarten were used unconditionally for multiple purposes with a restriction on a selection of supermarkets. Since this is a wide choice of supermarkets on the island of Sint Maarten, this modality has some similarities with the totally unrestricted modality of cash transfers. The evouchers are therefore not entirely comparable with cash transfers, but these can share a lot of similarities which is useful when translating this research to other modalities.

The starting point of this research is that only implementing agencies are able to change policies which make the context where beneficiaries are being helped the most. Humanitarian aid organisations have limited resources for each CVA project. This will be reflected in the outcome of this research, since it is not possible to help everyone with an infinite amount of money as an easy solution. Another point of warning is to avoid an endless debate on the consequences of helping vulnerable people. This can turn into an ethical argument on what most effective altruist policies are beneficial to recipients.

1.4. Research design

This section will introduce the Data Science Life Cycle, a framework to assist in structuring research. Each step will be explained on their relevancy in this research. Combined with the necessary background knowledge, this chapter will merge all the steps in a Research Flow Diagram and in an overview of the structure of the report.

1.4.1. Data Science Life Cycle

This research will be carried out following a transformed framework of the Data Science Life Cycle. There are different variations on data science process diagrams, data science life cycles and similar variations used to solve business problems and to conduct academic research (O'Neil & Schutt, 2013; Stodden, 2020). Using a framework with clear steps will guide the research in an organized way. Since data science projects and research problem differ, a tailored approach has been taken. This research derived the following revised data science life cycle to help solving the research question:

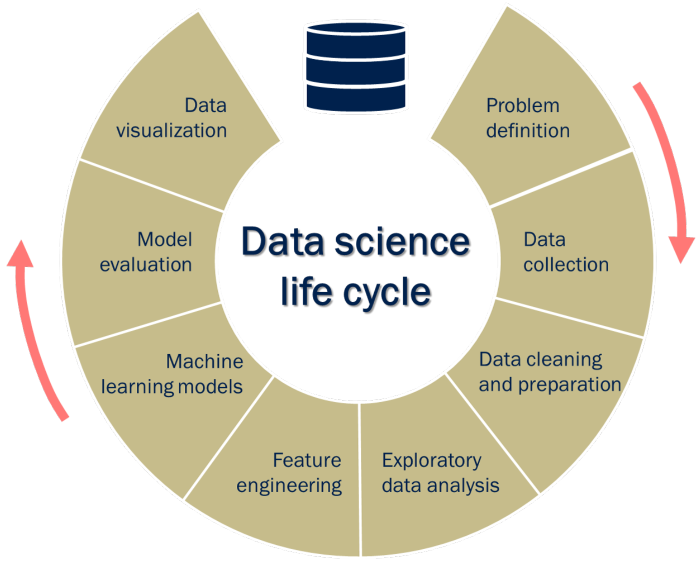


Figure 1.3: The revised Data Science Life Cycle

Problem Understanding

To solve any problem, the first step is to better understand the problem at hand. The introduction chapter already contributes to giving background information and by defining objectives of the research project. Further scoping is achieved by stating the main research question, sub questions and approach of this research.

Data collection

Data collection is an inevitable part of every data science project. Meaningful data is gathered via different sources. This data can be structured or unstructured and it should be clear how this data has been collected to know possible limitations.

Data cleaning and preparation

After collecting data, the cleaning and preparation process starts. Inconsistent and inaccurate data could lead to wrong and biased conclusions. This process involves deduplication, removing of missing values, correcting typos and formatting, removing blank spaces and correcting or removing inaccurate data.

Exploratory Data Analysis

When the data is cleaned, a first exploration can take place via Exploratory Data Analysis. With the help of statistical tools and multivariate visualization techniques a first insight into the underlying structure can be given. The objective is to understand and become familiar with the variables/features of the data before we can proceed to a more in-depth analysis.

Feature Engineering

Feature engineering involves processing and transforming variables to be able to use them in a meaningful way in machine learning models. This is done by enrichment of, and combination of existing variables using domain knowledge, such as knowledge on customer segmentation. Selecting the most important features by applying statistical analysis and pragmatic judgments is crucial in improving model performance. Feature engineering does not imply adding new data, but actually making existing data more useful for further analysis. For example, with the use of theories from customer segmentation, data from buying behaviour is transformed to an informative value, such as maximum amount spend or time between transactions.

Machine learning models

Machine learning models are selected for this research because 1) the world is complex and multidimensional and best understood by multivariate analysis and 2) we want to extract information from large datasets. Machine Learning (ML) algorithms are divided into four categories: supervised (with labelled data), unsupervised (with unlabelled data), semi-supervised (small fraction of data is labelled for training), and reinforcement learning (with unlabelled data which will be labelled) (O'Neil & Schutt, 2013). The research goals will influence the selection of specific ML methods, such as to classify, clustering, or predict variables. In this research, a clustering algorithm is used to extract segments of recipients of CVA. Because there are a lot of available variables, the clustering will be accompanied by dimensionality reduction methods (e.g., factor analysis and PCA) of which the motivation will be explained in chapter 6.

Dimensionality Reduction Since the data used in this research consists of many features, the dataset will be reduced on variable/column level. An exploration of factor analysis and principal component analysis is used to find the most (and least) explaining features. With these findings, existing variables can be removed or combined into new variables.

Clustering Clustering is unsupervised ML method we use when we don't know the label yet and where we should name the output ourselves. Clustering is an interpretable way of unsupervised learning. It is different from other methods since researchers do not have a certain assumption beforehand and the goal is not to create the best fitting clusters. In this case the goal is to better understand recipients of CVA, and clustering is a valuable method for achieving this goal. Since the evaluation of clustering is subjective, we need domain knowledge and expert validation to find useful clusters. We only know if clustering is appropriate if we can gather new insights.

Model Evaluation

Evaluating the machine learning models results is a crucial step to validate and check the fitness of the model for the data at hand. Different evaluation metrics, or so-called validation indices, are used to quantify model performance. These can also be used to compare different simulations of models. Validation indices of clustering include, for example, the Silhouette Index, Dunn Index and Gamma Index.

Data Visualization

As last step is visualising data used to explain results and translate them into clear insights. Telling the story from data with the use of charts, graphs and tables will be done with data visualization. Without this part, it can be hard to understand results of the analysis and therefore to implement findings into the real world.

Iterations

This Data Science Life Cycle has more iterations than the one shown in figure 1.3. As research progresses, many more sub iterations are possible. For example, when found during the modelling that there is not enough data available, one can choose to go back to the data collection step. The possible smaller iterations are not visualised in figure 1.3, since they will be explored during the research process.

1.4.2. Necessary Background Knowledge

Since this research aims to gain a deeper insight into recipients of CVA using cluster analysis, there is background knowledge needed. This necessary background knowledge is part of domain expertise on cash and voucher assistance, the case of Sint Maarten and customer segmentation.

Cash and Voucher Assistance

CVA exists for a long time, but has really taken off in last years. There are multiple modalities, mechanisms and other policies in the field of CVA which are important to understand in order to analyse recipients of cash and voucher assistance. Also targeting methods and project cycles are useful to further examine in order to successfully implement recommendations into new CVA projects.

Case selection: Sint Maarten

This research is applicable on a case which has been chosen partly because of its data availability. Background knowledge on the situation in Sint Maarten is important to put findings into context. The background story behind the data is as important as the data itself when stating conclusions in this research. Therefore, a chapter should be dedicated to the case study.

Customer Segmentation

There are multiple theories on customer segmentation which can be used in this research. The literature and applications should be explored further in order to use them for clustering recipients of CVA. Since marketing theories on customer segmentation are not used before, a specific portion of this research is reserved for extracting the most useful methods from this field.

1.4.3. Report Structure and Research Flow Diagram

The following combined diagram of the report structure and research flow diagram in figure 1.4 preserves the Data Science Life Cycle in different chapters and sections of chapters.

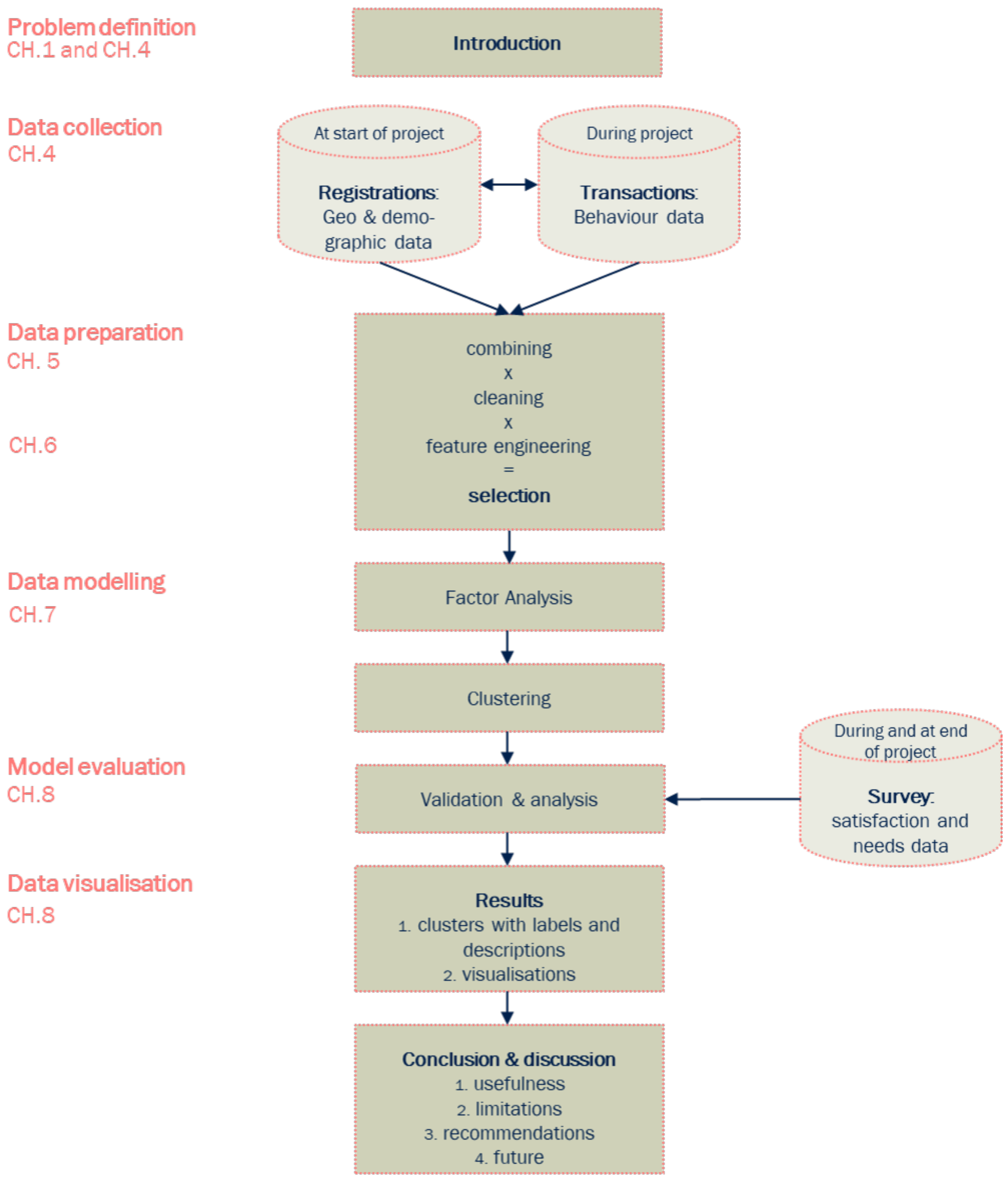


Figure 1.4: Research flow diagram with corresponding report chapters

2

Cash and Voucher Assistance

This chapter defines the core concepts and current state of research on cash and voucher assistance. Also, the project flow, commonly used indicators and policies of CVA are being presented. Next to explaining the implementation of CVA, aspects on privacy of recipients are also stated in this chapter.

The following subquestions are being answered in this chapter:

- What are common CVA configurations?
- How are CVA projects carried out?
- What are frequently used policies to target and help people in need?

In order to understand the workings of cash and voucher assistance, a literature review has been conducted. Search machines used are Google Scholar, Scopus and the online libraries of The Cash Learning Partnership (CaLP), The Cash Hub by The British Red Cross and the Active Learning Network for Accountability and Performance (ALNAP) for finding a great amount of relevant reports and papers. For findings related publications to an already relevant paper, the website of Connected Papers was used (Connected Papers, 2022).

2.1. Definition of core concepts

The term Cash and Voucher Assistance has related terms which are often used by different organisations. For example, the terms Cash Based Assistance, Cash Transfer Programming and Cash Based Intervention can all be used to describe roughly the same type of assistance (The Cash Learning Partnership, 2018). According to The Cash Learning Partnership, a collective of relevant actors in the field, the definition of CVA is the following:

> "CVA refers to all programs where cash transfers or vouchers for goods or services are directly provided to recipients. In the context of humanitarian assistance, the term is used to refer to the provision of cash transfers or vouchers given to individuals, household or community recipients; not to governments or other state actors. This excludes remittances and microfinance in humanitarian interventions (although microfinance and money transfer institutions may be used for the actual delivery of cash)."

There are different ways to implement a CVA program. First of all, the objectives of the program should be clear in advance. CVA could be used for covering basic needs, for a specific sector such as WASH (Water, Sanitation and Hygiene) or for multiple purposes. CVA can be given conditional or unconditional. With conditional transfers, there are prerequisites in order to receive assistance for example Cash for Work. Other requirements on CVA could be on the utilisation of CVA which make the assistance a restricted or an unrestricted transfer (Cash Learning Partnership, 2020a). Unrestricted CVA is always in the form of direct cash and restrictions are often implemented by using vouchers. With restricted CVA, humanitarian aid organisations can steer consumption of CVA towards a certain sector specific objective. The form of assistance can be two different categories: cash transfers and vouchers. Subcategories of vouchers include commodity vouchers where people can exchange a voucher for a fixed quantity of goods at participating shopkeepers. Value vouchers are different since they have a denominated cash value which can be used on multiple products. The most used delivery mechanisms of CVA are e-cash (often digital money sent to mobile phone wallets or bank transfers), physical cash in hand, paper vouchers and evoucher (a voucher stored digitally). The above-named types of CVA are important to highlight since most research focuses on a particular set of choices in a case and the workings of some combinations can work differently in different contexts.

In the project of Sint Maarten, the objective was to offer a multipurpose unconditional cash transfer. This has been done via restricted value vouchers which were delivered via evouchers. More details on the setup of this project can be found in chapter 3. An overview of the different possibilities on CVA programs can be found in figure 2.1. This figure can be seen as a summarized guide on the compositions of a CVA program. One can formulate, for example, a CVA project focused on basic needs, with unconditional qualifications and unrestricted access to beneficiaries. This example could be accompanied by the modality value voucher which is delivered via evouchers. Many more compositions are possible and highly dependent on specific context and needs of people.

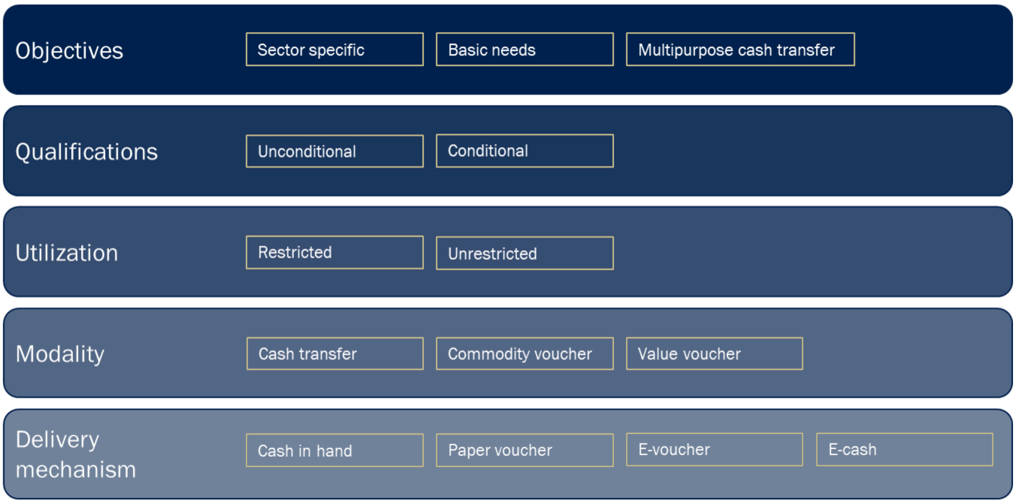


Figure 2.1: An overview of the different compositions of CVA

2.2. Project flow

CVA projects can be implemented by a certain sequence of steps. This section dives into the existing project flows of a CVA project and explores the addition of a feedback loop. A feedback loop will place this research in context by explaining how including behaviour data and users voices could be implemented.

2.2.1. Existing project flow

Most humanitarian aid organisations conduct CVA projects following a predefined framework or project flow. Although there exists multiple different guides on CVA project flows, there are many overlapping elements. The Cash Learning Partnership (CaLP) and the International Federation of Red Cross and Red Crescent Societies (IFRC) are two of the influential institutions on the design of CVA. By combining their project flows with elements from guides of other organisations, a summarized project flow of CVA projects has been constructed. This gives a clear overview of common practises of implementing CVA projects. The project flow can be seen in figure 2.2 and is based on various guides for CVA practitioners (Cash Learning Partnership, 2020a; ECHO, 2013; Harvey & Bailey, 2011; HelpAge International, 2010; International Red Cross and Red Crescent Movement, 2022; Sphere Project, 2011). Not all projects follow such a structured approach, but many will contain a large portion of elements presented.

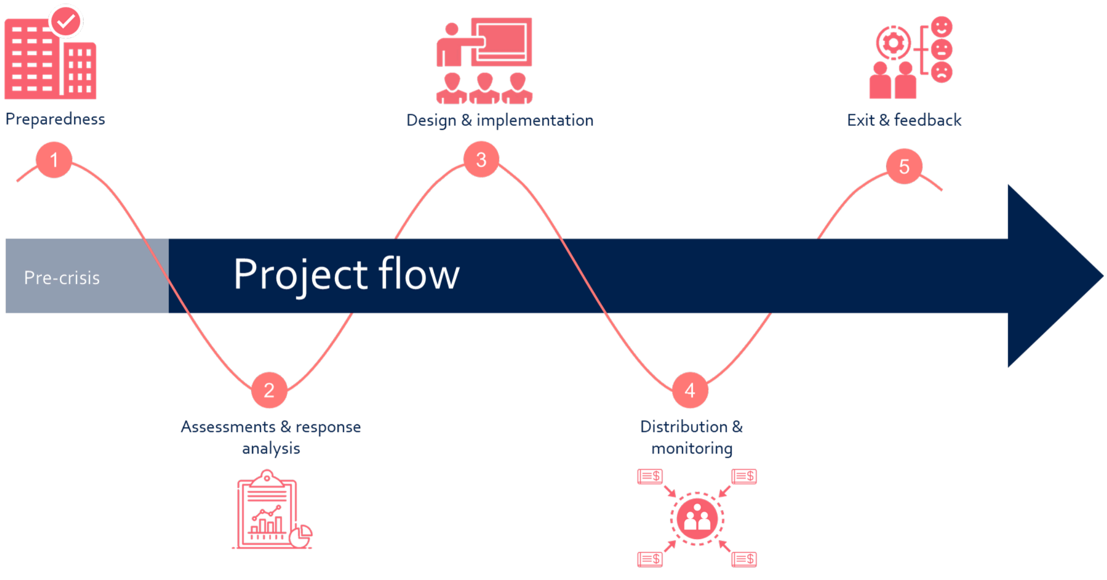


Figure 2.2: Project flow of CVA projects

A list of detailed aspects per phase will further increase understanding in how CVA project progress over time:

1. Preparedness
   - Baseline assessments
   - Contingency planning
   - Ready-to-go solutions
2. Assessments & response analysis
   - Cash Feasibility check
   - Needs and market assessment
   - Analysis of cash transfer value
3. Design & implementation
   - Plan of action (targeting and registration)
   - Contract suppliers
   - Standard operating procedures

- Implementation set-up (e.g., training of staff)

4. Distribution & monitoring

- Distribution
- Monitoring

5. Exit & feedback

- Post distribution monitoring
- Reporting & documentation
- Lessons learned

The first phase in the CVA project flow is already relevant before a crisis starts. Preparedness is regularly overlooked, but can have significant impact on the effectiveness of CVA projects (Van den Homberg et al., 2019). Contingency planning and having relevant systems ready for implementation are useful aspects of a high degree of preparedness. A good baseline assessment gives us information on the level of wealth and welfare of different subgroups of a country before the population is affected. Baselines are immensely useful since the goal of humanitarian aid is to bring the affected population back to pre-crisis levels. This is one of the main differences with development aid where the objective is to 'develop' affected population or economy to higher levels than pre-crisis (Lie, 2017). This is not always clearly separated in development and aid projects. Clear baselines are oftentimes absent in many countries where humanitarian aid is needed, making it very hard, if not impossible, to determine the baseline during pre-crisis times (Yeh et al., 2020).

When disaster (natural or man-made) strikes, various assessments will be carried out on needs, markets and on the feasibility of running a CVA project. The outcome of these assessments will determine if CVA is appropriate to use in the given environment. Different other analyses will conclude on potential cash transfers values and their expected impact on the observed needs.

The third phase is initiated when it has been found that CVA is an appropriate form of aid for the affect population. A plan of action will be constructed including details on how to reach and target beneficiaries. Details on existing targeting methods are discussed in next section. Registration and verification procedures are important for considering every eligible person or household. Since CVA projects often operates with the service of third parties (e.g., financial service suppliers), contracts will be signed with relevant partners and standard operating procedures are listed.

During the distribution and monitoring phase, the cash or vouchers are allocated among the qualified beneficiaries. Simple monitor mechanisms check if transactions were successful and if supermarket prices did not spike, due to having a small chance of inflation during CVA projects. From interviews with employees from the IFRC and from the literature (Obrecht, 2017), it was found that monitoring and collecting feedback is often overlooked in CVA projects, mostly due to time constraints. The distribution and monitoring phase can continue for months or years and anything in between. The decision on discontinuing a CVA project is made during this phase and can be the result of a variety of reasons (e.g., limited resources or a takeover by the local government). Important to know is that CVA projects often do not have a predefined ending date since it is unknown how quickly recipients will be relieved from their situation. It is not uncommon that projects are unexpectedly being prolonged for multiple times and that the transfer value and beneficiary selection also are reconsidered during this phase.

When the decision of ending a CVA project has been made, the exit and feedback phase starts. Oftentimes a post-distribution monitoring survey is conducted on a subset of the population to collect their opinions on this project. The aid organisations will write down lessons learned and document considered choices. An independent tax office and evaluation agency will examine the CVA projects on numerous targets suggested by donors of the aid organisations. This post-distribution monitoring survey can be enhanced by adding triangulation questions when asking data on satisfaction and needs. Control questions will add reliability of answers instead of collecting overly favourable answers.

Project flow with feedback loop

As introduced in section 1, this research explores an additional way of targeting and monitoring recipients of CVA. This can be made clear by proposing a novel feedback loop in the project flow. Data on voucher consumption can be collected when a CVA project has been running for a certain time. Also, data on satisfaction and needs could be collected by surveys (such as post-distribution monitoring surveys) in a sample of all recipients during the running of a CVA project.

Initial targeting uses geo-demographic data and needs of people to target and verify their eligibility to a CVA program. A relatively simple vulnerability or score from proxy means tests for eligibility of beneficiaries is used based on the geo-demographic data. When the total group of beneficiaries is known, an allocation mechanism decides the transfer value and assigned aid modality for each beneficiary.

When these vouchers are eventually consumed, and when survey data on needs and satisfaction has been collected, a feedback loop to a new allocation mechanism can be constructed. Based on new insights in recipients of CVA using behaviour data and intermittent survey data a smarter allocation mechanism can be used for next phases of CVA projects. The idea behind this new feedback loop is that with more insights in needs and consumption, people can be assisted more effectively, and limited resources could be more efficiently divided amongst the people most in need. Specific examples of additional assistance are explained in the next section. The feedback loop initiates a new targeting phase, which can be better called a 'retargeting' phase, of recipients to facilitate a smarter allocation of aid.

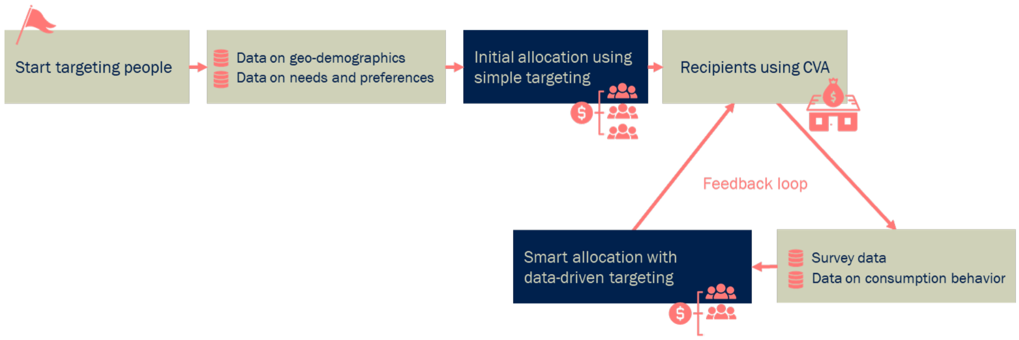


Figure 2.3: Feedback loop in new targeting flow

The limitation of this proposed feedback loop is that it operates with consumption and survey data, and these are only collected when the project is already running. Therefore, one could say that the new feedback loop suffers from a cold start, because it needs initial data to be used. The explanatory value of the data used in this feedback loop will be explained during this research. It is outside the scope of this research to come up with the specifics of a new allocation mechanism. First, research on the advantage of this new data is needed before new allocation mechanisms can be developed.

2.3. Policies and targeting

This section goes in detail in policies of how users of CVA can be assisted during CVA projects. Also, targeting methods are discussed to place the current targeting and added value of the feedback loop into perspective.

2.3.1. Targeting methods

Targeting methods of eligible beneficiaries include the following: self targeting, community targeting or proxy means test (PMTs) (Altındağ et al., 2021). Self targeting (or self-selection) is a method where people can consider themselves as eligible, while community targeting works with finding subgroups in the populations (e.g., people from certain geographical area) (Hanna & Olken, 2018). PMTs are frequently applied to target the poorest in humanitarian and development aid projects (Basurto et al., 2020). They use observable and verifiable household characteristics which act as predictor for household welfare (or similar indicators) (Sebastian et al., 2018). A PMT score is calculated and is used to rank households from poorest to wealthiest. This score serves as an eligibility threshold for participation in the aid project. In figure 2.4, we can see how using an eligibility threshold using a PMT score can create targeting errors. Since the PMT-score (on the y-axis) is a proxy for wealth, it is not perfectly equal to the real welfare of a household. The goal is to include every beneficiary under the poverty line, but by using a PMT-score we can experience inclusion and exclusion errors. The errors are not exclusively a result of using PMT-scores and can also emerge when using self- and community targeting.

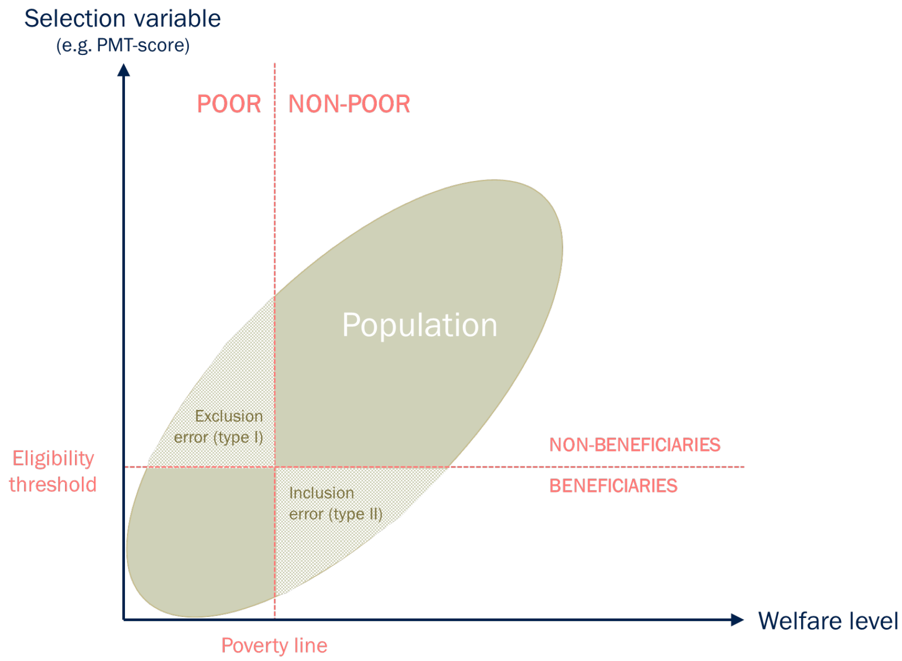


Figure 2.4: Visualization of targeting performance. Based on Sebastian et al., 2018

A limitation of the PMT is that it includes high in-built design errors. Literature shows that PMTs used in low- and middle-income countries only explain around half of the variation of poverty, so they only weakly predict the real situation of households (Kidd et al., 2017). Another constraint of PMTs is that they assume that when using simple survey questions, an accurate prediction of household poverty can be made. The reality is that it is more complex to capture a beneficiary in an often standard set of proxy variables (Kidd et al., 2017). This is another reason why this research could improve targeting (or retargeting) methods such as advancing PMT with measured data on consumption of aid combined with survey answers on needs.

2.3.2. Policies

Aid organisations have choices to make regarding the type of aid delivered. As already outlined in the first section, they have to choose the type of modality (e.g., in-kind, service, CVA etc.), distribution system (how and when to reach recipients) and delivery mechanism (e.g., evouchers, direct cash, paper voucher etc.). In reality, a combination of these could be meaningful to create synergistic impacts (Harvey & Pavanello, 2018). Also, the transfer value and frequency is highly important and usually correlates with higher satisfaction and psychological well being (Ground Truth Solutions, 2019). Next to this, choices on qualified products and rations are an efficient way to steer recipients into a specific kind of behaviour. For example, it is proven that nutritional diversity can be stimulated through CVA by only qualifying certain products for voucher assistance (CaLP, 2018).

The Netherlands Red Cross explains that choosing between decisions on CVA projects sometimes feels like guessing and trying and experimenting to see what works. There are multiple possibilities on long-term policies for CVA projects such as deciding on the scale and duration of the program, which are currently challenges for humanitarian aid workers (Dutch Relief Alliance, 2018). Effectiveness of CVA is strongly influenced by scale and duration of the program due to economies of scale (The World Bank Group, 2016). This effectiveness depends on the prices of goods that recipients can buy compared to the prices of goods aid organisations can buy, which can vary greatly.

Other ways to influence the impact on recipients are labelling and framing of cash transfers. A study of cash for education in Morocco showed that labelling the cash transfer for a specific purpose will be effective in reaching the desired outcome by the aid organisation (Benhassine et al., 2015). Other interventions during three cases in Tanzania, Kenya and Madagascar suggest an increase in impact by a little extra cost and effort (ideas42, 2019). These interventions were mostly based on educating and visualising the goals of the program for the recipients.

The use of accountability mechanisms, such as ways to collect feedback and complaints via whistle-blowing protocols, hotlines and via contact with stakeholders in the field is one of the policies which incorporate users feedback (UNFPA, 2020). This was also one of the recommendations on sensitive complaints mechanisms from the evaluation of the Sint Maarten project (Menéndez & Barrena, 2021). This type of community engagement can be an effective way to increase effectiveness of CVA while taking into account the voices of recipients. But it is a costly one due to the need of human capital in often remote areas.

If cash or vouchers during a CVA project do not reach the intended objectives and above policies are insufficient, other aid modalities too can always be considered afterwards. A broad selection of humanitarian aid could be to deliver food packages, shelter, hot meals, health services and aid related to water, sanitation and hygiene. Needless to say, this selection of other aid modalities should be dependent on the needs of beneficiaries and on capacities of the aid organisations. This ends this summary of additional policies which can be implemented during CVA projects.

2.4. Privacy and other aspects to consider

2.4.1. Privacy

There are two aspects which require special care when doing research in the humanitarian sector, especially when using data on payments and personal data. Critics in the humanitarian sector are warning of the risk of 'surveillance humanitarianism', 'data colonialism', 'digital exclusion', 'data injustice' and 'techno-solutionism' (Devidal, 2021). The point here is that digital payments, including evouchers, are not neutral and are often influenced by (political) objectives of partners such as banks, governments and tech companies. This could potentially jeopardise the operational independence of aid workers. The 'do no harm' principle is vital to organisations in the humanitarian sector, which also include to 'do no digital harm'. Vulnerable people receiving aid should have the same sufficient data protection rights. Also, with using a digital solution, they should not be discriminated against or excluded from programs. So, it will be important in this research to respect data protection rights and to include all people in need.

Data protection rights, according to the IFRC, follow responsibility principles for conducting CVA projects (2021). The principles are the following: 1) lawfulness, fairness and transparency, 2) purpose limitation, 3) data minimization 4) accuracy 5) storage limitation and 6) integrity and confidentiality. Especially purpose limitation and data minimization are important principles for this research since data has been collected which may not be used. The principle of data minimization states that one should collect as little as possible and only as much as necessary (IFRC, 2021). The purpose of the data collection should be made clear in advance. This forces practitioners to think about what (personal) data is strictly needed. Without these principles, there is a higher chance of collecting too much data and with this an increasingly higher impact of, for example, a data breach (510, 2018). Since this research dives into finding the most explaining behavioural variables of evouchers, data minimization can be further enforced in future projects since it will be known which data is not needed. For example, there could be data which is often collected in CVA programs, which does not have significant explanatory value. It will be beneficial for the recipient if these questions are not asked because it saves time and it minimizes the risk of digital harm.

Impartiality and neutrality are particularly relevant for data-driven projects in the case of the Red Cross and Red Crescent Movement. Impartiality means that is should not be possible with this data to discriminate against groups and individuals. Neutrality states that the motivation of analysing this data is not of political or economic interest. All reasonable measures must be made to ensure that the organization's basic beliefs and ideals are respected. Data must be acquired and utilised in a way that does not violate existing regulations or the organization's credibility. The data must be accessed and used in line with the law and the terms and conditions of third-party data suppliers. When planning to utilise data for any reason other than what is specified in their terms and conditions, explicit written consent from third-party data suppliers is essential.

2.4.2. Use of machine learning models in humanitarian sector

The use of machine learning (ML) models come with limitations, and especially the humanitarian sector should be aware of the risk of applying these models. ML models, and in broader term the field artificial intelligence, have a history of amplifying biases and acting as a black box when used in the wrong hands (Seltzer, 2006). Examples such as the "Toeslagenaffaire" in the Netherlands led to discrimination of vulnerable people based on using too difficult to understand algorithms (Klievink, 2021). Especially humanitarian aid projects focus on vulnerable people in need, who often have limited means to defend themselves to inappropriate use of ML algorithms. ML models can amplify implicit attitudes about people and can perform self-fulfilling predictions (Leavy et al., 2020). For example, real time algorithmic learning can bring different outcomes for two different subgroups of characteristics because of rarer or more common characteristics (Lambrecht & Tucker, 2020). Especially with vulnerable people from minorities, ML models can overlook their characteristics compared to more frequently available characteristics. This will be further elaborated in the discussion in section 8.

Another consequence of using data to measure people, and specifically in combination with machine learning models in the humanitarian sector, is the belief that everything can be measured and modelled. Measuring details of recipients in CVA projects can lead to thinking that one has a clear overview of what happens during such a project. But not every aspect of a recipient of CVA is possible to capture in a variable or in a ML model. There will be always gaps which are not being captured in data and these need extra care. Even with data-driven methods, humanitarian aid organisation still have to look for blind spots and act on them with a tailored approach.

2.4.3. Different views on CVA

Another important point to stretch is about the sometimes subtle difference between development aid and emergency humanitarianism. Aid for the development of poorer countries could be delivered in the form of cash and voucher assistance. This development aid is often politically incentivized and could be used to influence multilateral relations (Ark-Yıldırım & Smyrl, 2021). The purpose of humanitarianism is to "save life, alleviate suffering, and enable those suffering to maintain their human dignity during or after natural disasters and man-made crises" (Roger, 2007). This is also part of the core principles of the International Red Cross and Red Crescent Movement which is relevant because of the existing case study. The main goal here is to keep people alive and not to transform societies and economies. Although this transformation could be beneficial for people in the long term, the core principles of the The International Red Cross and Red Crescent Movement is to remain neutral and independent in order to help as many people as possible from suffering. Therefore, this research does not prioritise long term economic development or social transformation over helping vulnerable people, even though both could be achieved through cash and voucher assistance.

Then there is also the notion of paternalism in the humanitarian sector using a western perspective on aid and on what is 'good' for recipients. Most important of all is that we should not think for people in need, but we must involve them in the decision-making process. We can argue that it is ethically wrong to keep only account of, for example, the costs of a CVA project while not thinking about the factor of restoring personal dignity through CVA. This research analyses behaviour and satisfaction rates of people for analysing CVA projects. One of the starting points here is that the person should not change, but the context should be changed to help the person. CARE International and ideas42 offered this clear distinction in a study on improving CVA for women using behavioural science (CARE International & ideas42, 2020). But the reality is that there is a finite amount of money available which is earmarked for a most effective and efficient type of aid. That is why research on satisfaction rates and impact on recipients is so important to include users' voices and to avoid thinking for recipients.

3

Case Study: Sint Maarten

The case study of Sint Maarten will be illustrated in this chapter in order to be able to put results of the analysis of this research into context. The island of Sint Maarten with its economic drivers and dependencies will be described as well as the build up of different disasters which led to the evoucher project. Recent history of similar aid projects and evaluation of these projects are discussed along the reason why this case study is interesting.

The following subquestions are being answered in this chapter:

- What is the timeline of the CVA project on Sint Maarten and which actors are involved in the decision-making and setup of the project?
- What are the first interesting observations of this evoucher project and why is this a relevant case?
- What data is available for analysing recipients in our case study?

3.1. Introducing SXM

The island of Sint Maarten, often abbreviated to SXM, is part of two countries as can be seen in figure 3.2. The northern part is a French overseas collectivity called Saint-Martin while the southern side is a country in the Kingdom of the Netherlands. Sint Maarten is one of the most densely populated countries in the world with around 1.200 citizens per km2 with a total of almost 40,000 citizens (WorldData.info, 2022). Estimates of undocumented migrants range from 8.000 by UNICEF to a staggering 40.000 according to a controversial statement of the chief public prosecutor (Drayer, 2016; StMaartenNews.com, 2020; UNICEF, 2020).

The economy of Sint Maarten is heavily dependent on tourism since around 80 percent of the employed labour force is directly or indirectly related to the tourism sector (CIA, 2021). The many cruise lines docking in Philipsburg, the capital of Sint Maarten, are responsible for a big portion of the arriving tourists since 1.6 million people arrived per cruise ship and around 319.000 per airplane in 2019. With almost 48 tourists per capita in 2019, Sint Maarten ranks 4^{th} in the world for most tourists per capita. With tourism accounting for an average of 70 percent of GNP over the last decade, we can conclude that tourism is the driving factor on Sint Maarten (WorldData.info, 2019).

The labour force of Sint Maarten has a relatively high fraction of self-deployed and temporarily deployed working people. In 2017, around 20 percent of Sint Maarten was self-deployed compared to 17 percent in the Netherlands. Almost 25 percent of the labour force had a contract for temporary deployment in 2017, compared to 22 percent in the Netherlands and 11 percent in OECD countries (Hermans & Kösters, 2019). Around 6 % of the population is unemployed but taken into account the high estimates of undocumented workers, this percentage will be much higher. This makes the population relatively vulnerable to economic shocks.

Other characteristics of the island include the high dependency on import. Almost all food and energy is imported since the country has less means to produce its own goods, with the exception of local fishing and limited agriculture. In addition to the relatively high consumption by tourists and a low hourly wage of inhabitants, this results in a high cost of living for citizens of Sint Maarten (CIA, 2021).

3.2. Disasters and impact

To understand the shocks that Sint Maarten endured, a description of the two most recent natural disasters is presented. Hurricane Irma came ashore on the 6^{th} of September in 2017 as one of the strongest hurricanes ever recorded with wind speeds up to 296 km/h. This resulted in severe structural damage and in 4 deaths and 34 injured on the Dutch part of the island (Sterling, 2017). A geographic analysis accompanied with a survey conducted by the Red Cross and 510 found that around 90 percent of the structures on the island were damaged and one third of the buildings were destroyed (CBS News, 2017). The World Bank estimated the damages and losses of around 260 percent of the GDP of Sint Maarten (World Bank, 2021). Because of hurricane Irma, GDP dropped by 4.7 percent in 2017. Furthermore, in the last decades, tourism in Sint Maarten accounted for a yearly average of around 70 percent of total GNP which dropped to almost 40 percent of total GNP after the hurricane (WorldData.info, 2019).

In the aftermath of the hurricane, a grim situation of people looting and walking with weapons originated. French and Dutch army troops were sent in to control the situation. Together with the military, the Netherlands Red Cross went with two military planes full of in-kind aid to the island to assist in alleviating suffering of citizens. The King of the Netherlands Willem-Alexander visited Sint Maarten and was shocked by the destruction and immediately called for support from the European Union. The EU responded by allocating 2 billion euro in emergency funding to restore basic essentials (Darroch, 2017). Next to financial help from the EU, citizens of the Netherlands donated to the Red Cross Netherlands to raise money for aid relief.

It is believed that more hurricanes will become more frequent in these parts of the world because of rising sea levels and more extreme weather patterns due to climate change (IPCC, 2021). One of the natural factors mitigating the effect of hurricanes is seagrass beds which are important for anchoring sand and decreasing wave strength to prevent flooding (Royal Netherlands Institute for Sea Research, 2020). Sint Maarten experiences a significant decline in seagrass in the last 30 years and this will leave the island extra vulnerable for coming disasters (Nature Foundation Sint Maarten, 2016).

After a period of rebuilding the island and by multiple relief projects by the Red Cross, the well-being of citizens and the state of the economy of Sint Maarten steadily improved over the years. Until the first case of Covid-19 was registered on the 17^{th} March of 2020 (Voncken, 2020). Lockdown rules caused stay-over arrivals dropped to 438 arrivals in 2020 Q2 compared to 85,602 in 2019 Q2 resulting in an extreme change of 99.49 percent (Ministry of Economic Affairs & Telecommunications, 2020). Not all tourism was declining so rapidly during the Covid-19 crisis since cruise tourism went down by 16 percent in 2020. The total economy contracted by 24.8 percent in 2020, which could have been more without the liquidity support from the Netherlands of 13.3 percent of the monetary unions GDP (International Monetary Fund, 2021). Unemployment because of the Covid-19 crisis was estimated at 16.9 percent. Although the negative effects of this crisis on employment was probably even higher because of a perceived increase of underemployment (individuals who have a job, but they are not able to work as long as they would like to and/or the job is not up to their standards) due to absence of data (International Monetary Fund, 2021).

3.3. Setup of aid projects and evaluation

The Red Cross carried out multiple aid projects since 2017. It is relevant to know which projects were fulfilled before the evoucher project used in this research is presented. By knowing what kind of aid the recipients have already received in the past, we know their familiarity with different aid modalities and we can later place results in context.

3.3.1. Prior to the case study

After the hurricane in 2017, a National Recovery and Resilience Plan (NRRP) was written which existed of 4 phases in a 7-year period (Government of Sint Maarten, 2017). The first three months were dedicated to emergency response, the first year to immediate needs and early recovery, the second year to short-term needs and recovery and the rest of the 5 years on medium- to long-term needs and for rebuilding a resilient society. One of the reasons for a relatively slow recovery of Sint Maarten was the delayed funding to the NRRP of the government of Netherlands of around 600 million dollars (De Hamer, 2019). The Dutch government demanded two conditions to be met to the Sint Maarten government before setting up the fund. One of those criteria was on institutionalizing an anti-corruption body which should have the mandate to give binding advice to the government of Sint Maarten. Since this existing desire of the Dutch government has already been debated for many years, it took over seven months before a trust fund was constructed which unavoidably caused delays in executing rebuilding projects to the point that many houses still were not repaired 18 months after the hurricane. Another reason for slow recovery is the lack of coordination and abundance of aid organisations wanting to help according to the head of disaster management of Sint Maarten (De Hamer, 2019).

In addition to the NRRP, the Red Cross Netherlands organised a national campaign day just a week after the storm had passed resulting in a total of 5,247,863 euro. With these donations, the Red Cross Netherlands together with the Red Cross Sint Maarten was handing out emergency kits and necessary first aid in the first weeks. Distribution took place via different methods: via community-based organisations, via individual relief teams and by handing out to everyone who came to distribution points (Netherlands Red Cross, 2017). The Red Cross Netherlands helped around 26.701 people with emergency assistance in the first two months.

Since the 8th of November, among other aid modalities such as food packages and hot meals, food vouchers have been distributed. The goal of these food vouchers was to help the most vulnerable people (which were selected on income and household size) to do their own shopping at a selection of supermarkets. Since recipients were not allowed to buy tobacco or alcohol, these food vouchers could be seen as unconditional and semi-unrestricted ways of voucher assistance. This initial project had a duration of 2 months were 2.811 households where being helped with four vouchers worth 83.33 US dollar (Netherlands Red Cross, 2018). This initial project experienced some chaotic first distributions since the selection criteria were not made clear which created differing expectations. After some lessons learned of improving communication and digital registration processes, this project was extended for a number of times until august of 2018. In total 4,222 households received four vouchers worth of 83.33 US dollars in 10 months time (Netherlands Red Cross, 2019).

Figure 3.1: One of the first food vouchers on Sint Maarten (Soualiga News Today, 2017)

When Covid-19 entered Sint Maarten in March 2020, the Red Cross and other aid organisations were still working on aid and recovery projects from hurricane Irma, for example by handing out hot meals, building materials and food packages (Netherlands Red Cross, 2020b). The Netherlands Red Cross presented a new 'National plan of action COVID-19' to assist in the prolonged crisis on Sint Maarten (Netherlands Red Cross, 2021a). A new evoucher project started in May 2020 which had similarities with the previous voucher project. One of the differences is that this evoucher was an electronic debit card which received a monthly or two-weekly top-up instead of a one-time used paper voucher. The biggest change however was the renewed back-end of this project since the process was streamlined using several digital systems including a communication channel to reach all the recipients via SMS or WhatsApp (Meijer, 2021). Digitisation of this renewed evoucher project including assistance has been executed by 510. This project is the subject of this research since most data is collected and since this research is supervised by 510.

3.3.2. Case study

The evoucher project (together with providing meals and food packages) lasted from May 2020 to May 2021 and was financed by the Ministry of the Interior and Kingdom Relations with an initial budget of 16,000,000 euro for the islands of Aruba, Sint Maarten and Curacao. At the end of the project in May 2021, the project was transferred to the local government. A total of 3,145 recipients received a total of 5,653,687 US Dollar (approximately 4.7 million euro). The goal of this evoucher project is to support people to meet their basic needs.

Before this project started, a cash feasibility analysis has been conducted to get a new understanding of the context (Netherlands Red Cross, 2020a). The following aspects have been accomplishment in this analysis: needs assessment, stakeholder mapping, community assessment, market analysis, mapping infrastructure and financial service providers, organizational capacity analysis and possibilities on modalities and delivery mechanisms. The needs assessment estimated the following distribution of monthly expenditures for affected people: 50% for house rent, 15% for utilities, 25% for food and 10% for other expenditures. This aid project aimed at supporting the food and other expenditures categories. From the market analysis, five out of eight supermarkets were chosen as most suited for handling vouchers at strategic points on the island.

Four different options for delivery and modality were being considered: unconditional cash with checks or envelops as delivery mechanism, rechargeable evouchers to buy food or hygiene products at specific supermarkets, paper vouchers to buy food or hygiene products at specific supermarkets and in-kind distributions of food and hygiene products. Via a multi-criteria analysis the rechargeable cash evouchers supported by data and evoucher management platform Redrose were chosen as most suitable delivery mechanism. The proposed transfer amount was set on 85 USD for 1 to 3 household members, 128 USD for 4 to 6 household members and 183 USD for more than 6 members. Children under the age of 2 counted double for this allocation formula, since hygiene items for babies are relatively expensive. The initial connection of transfer amount with household size was later changed. The following households were immediately eligible: households with no member getting any kind of incomes, not receiving government support, not owning a house and with children under 5 years old. Once this selection had been approved and there was still room for more recipients in the project, households with pregnant women, single-headed household and big household were also eligible.

An evaluation report of this evoucher project pointed out that accessibility to people with disabilities could be improved and that the "registration system was not the most adequate to ensure that all vulnerable people had easy access to registration" (Menéndez & Barrena, 2021). Also, around 20 percent of the eligible people for this project did not prefer CVA over in-kind assistance. In this evaluation of this project interviews were conducted to explore if recipients perceived evouchers above unconditional cash. Several people from the interviews expressed their concern that these vouchers would be misused for purchasing alcohol, cigarettes or lottery tickets.

3.4. Relevance of this case

Using a specific case for evaluating CVA projects is a good way for researching on how to improve CVA. Especially people from the Caribbean are not being heard well enough according to the Cash Learning Partnership (CaLP) (Cash Learning Parthership, 2020). Compared to other global clusters, there is limited evidence for CVA in the Caribbean and this has "restricted the ability of evidence-based decision-making on the allocation of limited resources". Furthermore, CaLP stresses the need to evaluate projects to ensure that CVA is still effective in meeting the various and changing needs of affected populations. Analyses of past projects can be used for advocacy and to acquire new stakeholder buy-in (Cash Learning Parthership, 2020).

This CVA project on Sint Maarten is particularly interesting because there is large dataset available on vouchers and demographics of recipients. The Grand Bargain workstream states that granular accessible data on transfer values is still scarce and could be an improvement if made publicly accessible

(Development Initiatives, 2021b). There has been a project evaluation from a qualitative point of view but not yet from a quantitative view using this unique dataset. Furthermore, aid workers in Sint Maarten presume that some interesting emergent recipient behaviour happened such as saving, stockpiling and that there are differences in spending patterns in subgroups and characteristics of recipients. Also, the people on Sint Maarten had to deal with different kinds of disasters, such as an abrupt shock by hurricanes and lingering stresses from the current Covid-19 crisis. This makes Sint Maarten a very interesting case to use for researching CVA and to help decision-makers of CVA projects in the future with doing better quality CVA in different kinds of disasters.

3.5. Data Collection

The following data is available to use on the CVA case of Sint Maarten. The appendix D presents a more detailed overview of the usable variables.

Household data
Household data on demographic and socio-economic situation submitted by people via a survey in order to participate in the CVA project. This data also encompasses their preferences for aid modalities (CVA, in kind and hot meals). This dataset is anonymized and aggregated on neighbourhood level.

Voucher data
Voucher data on which household spend what amount of money at what time Supermarket receipts data on which product has been bought for what monetary value. This is possible to connect to voucher data. This data is geo-located and timestamped. Part of this is stored digitally and part of this is still physically available in a folder in the form of purchase receipts. This may be relatively easy to digitise with a coding script if it is desirable to expand the voucher data. This is also dependent on the possibility of receiving the receipts in time.

Post-distribution monitoring (PDM)
This consists of a satisfaction survey and includes both the demographic and socio-economic aspects of households. This data is not linked with earlier mentioned datasets, but it might be possible to make some connections on neighbourhood level.

Other possibilities that can be further explored is to use GIS data from the state of houses (broken or not) on Sint Maarten to compare this with the data on households and neighbourhoods. 510 has been mapping Sint Maarten quite elaborate via their mapathons. Secondary data such as economic indicators from the World Bank, GIS files from 510, and various statistics from the government. The following first interesting observations from the post-distribution monitoring survey in table 3.1 made this an initial interesting case.

Experiences	Percentage
Said that the voucher was not sufficient for their needs	9
Said that the voucher was somewhat sufficient in fulfilling their needs	40
Experienced increasing prices	48
Had suggestions for improvement	10
Did not know how to communicate or complain	16.5
Felt not safe	11
Experienced bad communication	3
Experienced bad quality distribution	2.6
Had to pay in order to get vouchers (sign of corruption)	1.4

Table 3.1: First observations from the PDM survey

3.6. Red Cross and 510

As mentioned before, this case is focused on one of the projects executed by The Netherlands Red Cross in close collaboration with 510 and the local Red Cross branch of Sint Maarten. To illustrate the goal of the CVA project of these actors, their mission and their relationship to each other has been stated.

The International Red Cross and Red Crescent Movement has 193 individual national societies of which the Netherlands Red Cross is one of. The Red Cross Sint Maarten is an overseas branch of the Netherlands Red Cross, but acts as a separate national society since it is registered under the law of Sint Maarten with an auxiliary role. The Red Cross Sint Maarten has the same objectives as the Red Cross Netherlands and they were of assistance during all aid projects on Sint Maarten. The mission of the Red Cross is to prevent and alleviating human suffering everywhere, protecting life and health and ensuring respect for people with special attention to those who are most vulnerable during armed conflicts, disasters and other emergencies (Netherlands Red Cross, 2021b).

510 is an initiative of the Netherlands Red Cross and in this evoucher project, they were responsible for the digitalization of vouchers using several digital tools. Their mission is to "Shape the future of humanitarian aid by converting data into understanding, and put it in the hands of humanitarian aid workers, decision-makers and people affected, so that they can better prepare for and cope with disasters and crises" (510, 2016). Furthermore, 510 aims to achieve this goal by smart use of big data for faster and more (cost-)effective humanitarian aid and by increasing understanding in humanitarian data.

Figure 3.2: Sint Maarten Districts (UN OCHA, 2019)

4

Data Preparation

This chapter on data preparation goes more into detail in the data cleaning methods and which choices are made during this process. A motivation of the use of different programming languages is given and a clear overview of the data flows is presented.

The following subquestions are being answered in this chapter:

- How will the datasets be cleaned and combined?
- What does exploratory data analysis tell us about this dataset?

4.1. Data Flow Diagram

The datasets are cleaned in various notebooks and exported to other notebooks and scripts for other operations. This research makes use of Jupyter Notebooks based on the Python programming language and of scripts in RStudio based on R. To visualise the full journey of the data in notebooks, a data flow diagram has been created which can be seen in figure A.1 in appendix A. The rectangles and circles are respectively Jupyter Notebooks and R scripts, while the cylinders on top are the original sources of data. Each flow has been accompanied by a specific name of mostly CSV-files which are imported or exported. There are in total 5 different sources of data which can be summarized in two different flows: three sources are combined from the datasets on demographics while two other sources are combined from the data of the survey. The Jupyter Notebooks and scripts in R are accompanied with this thesis report and can be seen on Github. The following more elaborate description of the data preparation will be presented with this data flow diagram in mind.

4.2. Data Cleaning

Methods used and notable choices during the data cleaning process are described. Also, the treatment and potential value of the survey data is discussed.

4.2.1. Methods

The goal of data cleaning is to remove or fix any data which is incorrect, inaccurate, incomplete, incorrectly formatted or duplicated. The detailed process of cleaning these datasets can be read in the notebooks and a global overview of the cleaning process is presented in this paragraph. To get an overview of all the original variables, appendix D elaborates more on the specific variables and their descriptions.

The first goal is to merge datasets on related ID variables so that we get the full picture of all the variables of all recipients. From there on, the dataset is trimmed. The data on demographics (BNF nonPII and HHmembers nonPII) has been successfully merged so that a dataframe of 6,371 persons by 63 variables is created. Since this data also contains recipients of food packages and hot meals, these entries will be removed. An inspection of expected datatypes showed some entries with numerical values which should have been non-numerical and vice versa. Changing datatypes and inspecting wrong entries gives already a good overview of how this data should be cleaned. There are multiple values which have been spelled differently but have the same meaning. Minor transformations such as changing capital letters made the data more standardized. By the use of plotting histograms, boxplots, z-scores and correlation plots, certain outliers have been detected and corrected or removed according to common sense. Variables related to income are standardized to the currency of USD using a generally accepted fixed currency rate of 1 ANG to 1.79 USD. Variables with a high percentage of missing values have been removed since they will not be useful in the machine learning models because of too small sample sizes. For cleaning the data from the survey, the process is very similar. The demographic data from the registration was reduced to a size of 5,491 recipients with 42 variables while the survey data is reduced from 1,749 recipients on 145 variables to 754 recipients by 73 variables. The explanation for this high reduction of rows and columns is that these sources also included data on food boxes and hot meals, two of the other aid modalities from the projects of the Red Cross on Sint Maarten. This data is not relevant for this research and is therefore removed.

Cleaning voucher data (NLRC Vouchers) is more time-consuming, since the original dataset has the size of 184,404 rows by 13 columns. This dataset consists of rows on the amount of USD spent by recipients, received by supermarket and transferred by the NLRC. The data of interest is data regarding buying behaviour of specific recipients, so we are only interested in money spent by recipients and money transferred to recipients by the NLRC. Money received by supermarkets is removed from this dataset because of irrelevancy. The dataset is further cleaned by using boxplots, z-scores and interquartile range measures. The goal is to transform this dataset from a time-series with expenditures on the x-axis (rows) to unique recipients on the x-axis (rows). With this format, merging with the other datasets can take place using the recipient IDs. With the use of feature engineering, which can be read in chapter 5, new features have been constructed with this expenditure data.

After initial cleaning of all raw datasets, some datasets were combined even further. The cleaned geo-demographic data from the registration, and the transformed voucher data are combined to a new dataset (df_combined.csv). During this last merging of datasets, even more recipients were removed from the registration data because they did not participate in the voucher program. This led to a final size of the dataset of 2,888 recipients with 66 variables.

4.2.2. Merging data with survey data

This research could benefit from having all the available data of every recipient ready for clustering in one dataset. Data on needs and satisfaction could be useful for directly observing problems of recipients during projects. Unfortunately, the survey data is being collected anonymously in this specific case, so there are no IDs to merge datasets on. It has been tried to find the overlapping perfect pairs in the survey and demographic datasets, but this can not be done with sufficient accuracy. The survey data was collected approximately a year after the registration data has been collected and this causes different results for household compositions which can not be traced back. For example, we only know the amount of people between 3 and 17 years hold in a household, but we don't know if someone is of the age of 17 in this composition and if this person has been transferred to the household composition category of people with the age between 18 to 65 years a year later. This also applies to other household compositions and makes it inaccurate to securely match both datasets on overlapping variables. Therefore, another approach has been used which compares results from the survey data with the found cluster results. For future projects, it can be decided in advance to connect registration data with survey data.

4.2.3. Choices made

During the cleaning process some choices had to be made which can influence the outcome. One choice which applies to the demographic and survey data is about answers where people could fill in whatever they deemed necessary. These answers range from a few words to sentences in multiple languages. Most questions were about recommendations to the NLRC, an estimation of the items they bought and about descriptions of their medical situation. To get an idea of the answers of these variables, a Word Cloud has been constructed which visually represent the relative importance of text data. It was found that this data is too hard to transform into usable variables than can be analysed using cluster analysis, so this data has been removed.

Another choice has been made to not include specific medical data. Recipients had the choice to tell their specific medical problems. Since this is detailed sensitive data, it was chosen to not include these variables in the dataset. An exception has been made for the number of disabled persons per household since this is a specific group of recipients which had more problems according to the evaluation done by an independent agency in 2021 (Menéndez & Barrena, 2021).

A choice worth mentioning in cleaning the voucher data included the removing of negative values. Exactly 24,892.10 USD has been paid back by recipients to the Red Cross. This money could have been refunded due to many reasons such as recipients being found ineligible after they received the vouchers. Another reason is possibly that at the end of the project, all the accounts were being balanced and any leftovers are transferred to the Red Cross again. Since this is only a small percentage (0.22 percent) of the total vouchers spent, this is negligible to further examine and has been removed.

4.3. Exploratory Data Analysis

The cleaned datasets are further explored to get a clear overview of the story behind the variables. Where possible, the data is compared by available data from the Government of Sint Maarten to find similarities. Exploratory data analysis is important to quickly test assumptions and to discover trends and underlying structures otherwise not easily observed.

4.3.1. Comparison with government data

The Department of Statistics of Sint Maarten published two interesting reports on census data, among other data. This is useful for comparing the registration and survey with the data from the government. This makes it able to see potential differences between the three datasets and these comparisons have been done with data on age, income and district. This is as far as we can come with comparing the research data with other publicly available data.

Age

To compare the age distribution of the Department of Statistics with the age of household members we can not use one diagram. This is because the census data has an interval of 4 years and the household composition data has deviating intervals which can not be directly compared (Department of Statistics Sint Maarten, 2017). Therefore, a histogram of the population from the census data is presented in figure 4.1 while the registration and survey data on age is presented in a table 4.1.

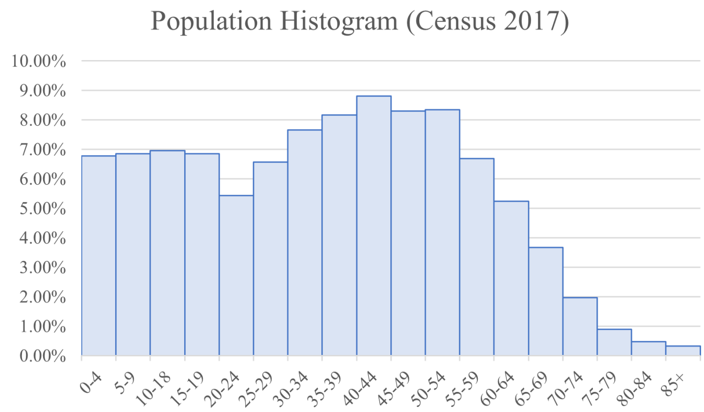


Figure 4.1: Histogram of Population of Sint Maarten (Department of Statistics Sint Maarten, 2017)

Number of household members with age	Registration	Survey
0-2	5%	9%
3-17	27%	28%
18-64	62%	56%
65+	7%	7%

Table 4.1: Household composition of registration and survey data

It can be concluded that the registration and survey data are much alike, when looking at household age compositions. There is a similar number of elderly people above 65 years old in the census data (around 7 percent). Another conclusion can be made regarding young people under 18 years. There is a relatively high number of younger people in the CVA project compared to the census data. Therefore, the recipients of CVA consists of families with a relatively high number of children compared to the census data. An assumption here is that the population distribution did not change from 2017 to 2021, since the census data is from 2017 and the registration data from 2021.

Gross Monthly Income per Household
Another comparison can be made with household income between the census data from 2018 and registration data from 2021 (Department of Statistics Sint Maarten, 2019). A table with household income in intervals of 1,000 ANG has been constructed to check for similarities in both datasets. For comparison purposes, also the equivalent in USD is noted.

Monthly income in ANG	Census 2018	Registration
No income	5.0%	24.5%
1 to 1,000 (559 USD)	14.0%	45.9%
1,001 to 2,000 (1,117 USD)	22.0%	27.8%
2,001 to 3,000 (1,676 USD)	14.0%	1.4%
3,001 to 4,000 (2,235 USD)	13.0%	0.2%
4,001 to 5,000 (2,793 USD)	7.0%	0.1%
5,001 to 6,000 (3,352 USD)	6.0%	0.0%
6,001 to 7,000 (3,911 USD)	5.0%	0.0%
7,001 to 8,000 (4,469 USD)	3.0%	0.0%
8,001 to 9,000 (5,028 USD)	2.0%	0.0%
9,001 to 10,000 (5,587 USD)	1.0%	0.0%
10,000+	10.0%	0.0%

Table 4.2: Income frequency of Census 2018 and Registration data

It can be concluded that the registration data includes the bottom of the household incomes of the population of Sint Maarten. The assumption with this conclusion is that household income did not change drastically between 2018 and 2021. The household income from the census was collected after hurricane Irma struck Sint Maarten and before Covid-19 entered the island. It can be argued that household income did not change much after the hurricane to the beginning of the registrations, since most jobs were already ended due to a high dependency on tourism.

District
The number of residents per district is also collected during the census of 2017 and during the collection of registration data (Department of Statistics Sint Maarten, 2017). There are 8 districts which locations can be seen in figure 3.2.

District	Census 2017	Registration
Colebay	17.7%	17.6%
Cul-de-sac	21.2%	19.6%
Little Bay	13.8%	9.7%
Low Lands	1.7%	0.2%
Lower Princess Quarter	26.7%	30.2%
Philipsburg	4.7%	7.4%
Simpson Bay	2.8%	3.0%
Upper Princess Quarter	11.3%	12.2%

Table 4.3: Residents per district of Census 2017 and Registration data

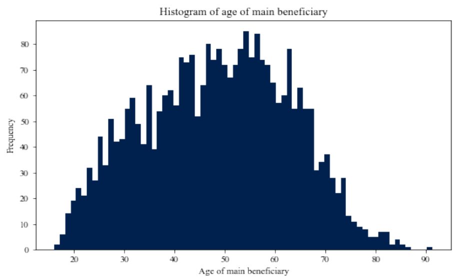
The residents per district in both census data as registration data is quite similar. The districts Low Lands and Cul-de-sac are underrepresented in the CVA project, while the districts Lower Princess Quarter and Philipsburg are overrepresented. This can be clarified with the difference in income per district. It must be noted that this conclusion assumes that there were no significant internal migrations to other districts in the years from 2017 to 2021.

4.3.2. Registration and Survey data

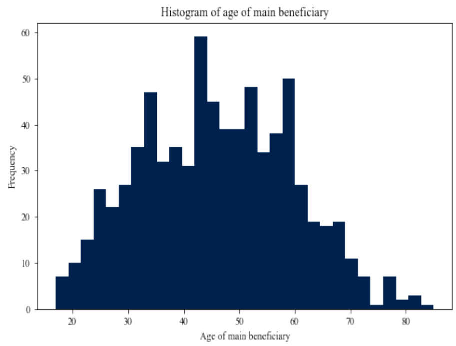
To further explore the registration and survey data, the overlapping variable on age is further investigated and other interesting variables on financial situations are inspected. The registration and survey data include many open questions where multiple divergent answers in numerous languages are given. These questions have been excluded from the analysis, although they could add value to the analysis. Future CVA projects should minimize the use of open questions and only include them when really necessary and when closed questions are not sufficient.

Age of main beneficiary

In both the survey as the registration data, the age of main beneficiary has been asked. The histogram of age in both datasets is presented in 4.2. Since the mean, standard deviation and distribution are very similar, we have an indication that the survey data is indeed a sample of the registration data. A two-sample Kolmogorov-Smirnov test has been used to check if the two datasets come from the same distribution. It was found that the variable age, among other overlapping variables, are samples from the same distribution.



(a) Registration data



(b) Survey data

Figure 4.2: Histogram of the age of the main beneficiary

Financial situation

The financial situation of the household has been collected in the registration data. Four questions were asked on the amount of: *AOV* (government pension), *Onderstand* (social assistance from the government), *SSRP* (unemployment support) and their monthly salary. To explore the relation between those four sources of income, a bivariate pairwise scatterplot is displayed in 4.4d. The diagonals are used for the distribution of the specific source of income. We initially find that there are no visibly interesting relationships between, indicating that households have only one main source of income instead of having multiple sources.

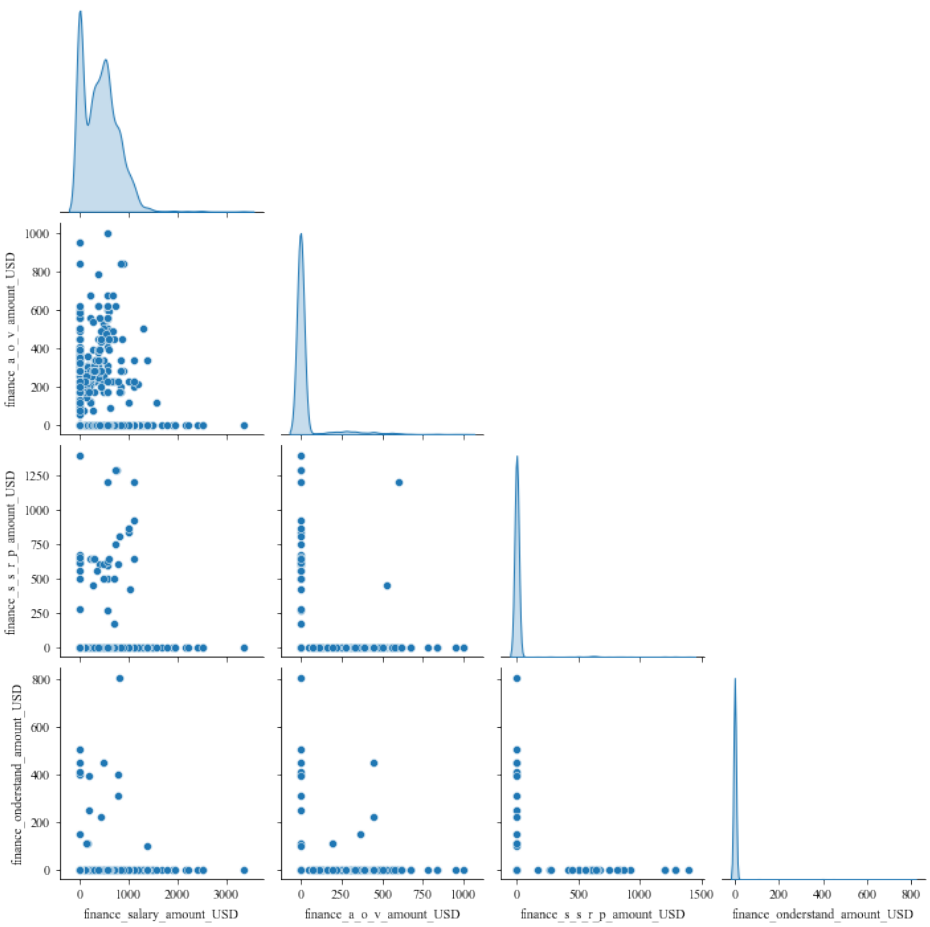
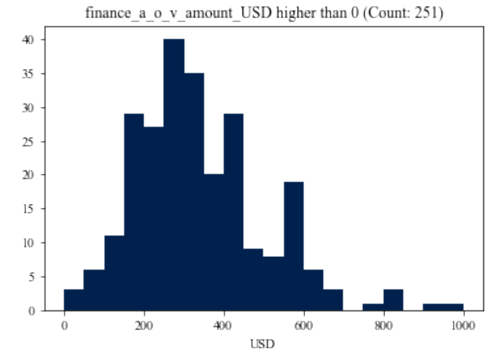
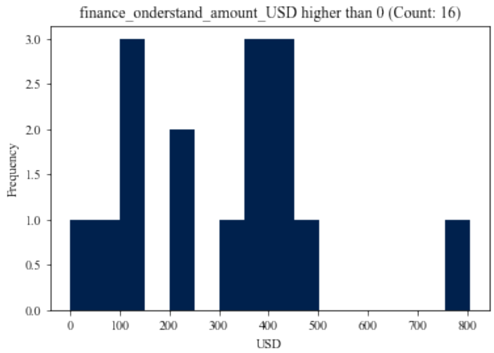


Figure 4.3: Pairplot of financial situation

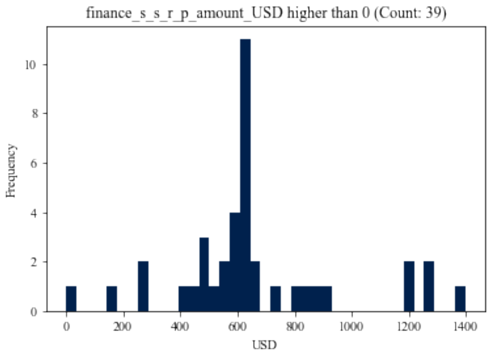
The distribution of the financial situation variables are heavily right skewed due to a large number of people receiving nothing from that specific source of income. Therefore, we also visualize the distribution of each source of income excluding households who don't receive (i.e., specific source of income above zero). This gives a bigger understanding in how households cope without the evouchers. Only a small percentage of the households receive *Onderstand* and *SSRP* while around 9 percent (251) of the households receive some form of pension.



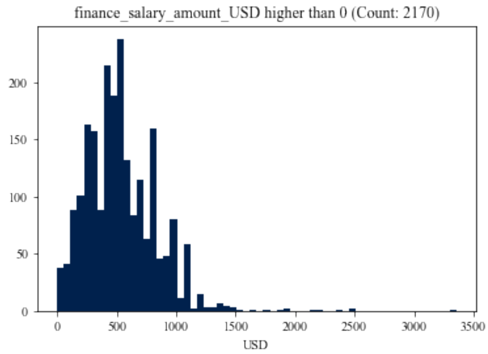
(a) Distribution of monthly pension



(b) Distribution of monthly social assistance



(c) Distribution of monthly unemployment support



(d) Distribution of monthly salary

Figure 4.4: Financial situation distribution

4.3.3. Voucher data

For visually exploring the data of evouchers, the amount spent by recipients has been plotted over the duration of the project (one year) in figure 4.5. It is clear that the project started with only a few visits to the supermarket in the first months. The biggest sum of vouchers have been spent in the last half year of the project. Another observation is that the vouchers have not been spent equally over time, but have been spent at concentrated days in time. The first 6 months of the project, the vouchers were spent once in two weeks, while in the last 6 months vouchers where spent monthly. This can be explained by the change of policy to distribute vouchers less frequently in the last 6 months.

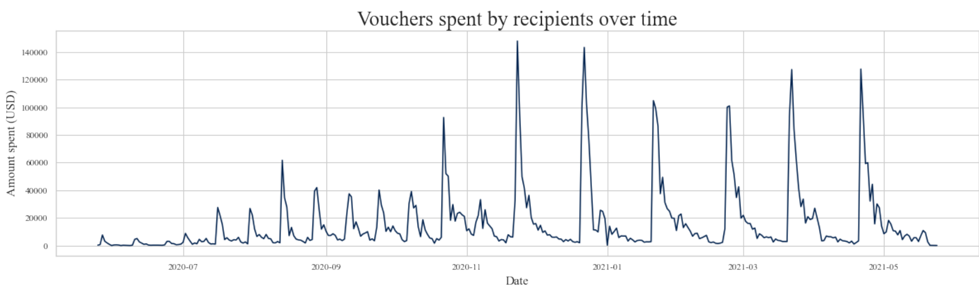


Figure 4.5: All vouchers spent by recipients over time

To further examine the spikes in the voucher expenditures, a connection has been made with the sum of vouchers received (in USD). Figure 4.6 shows that there is a short time between the money is received by recipients and when it has been spent. We can identify an equally right skewed distribution of money spent on vouchers after each time vouchers have been received.

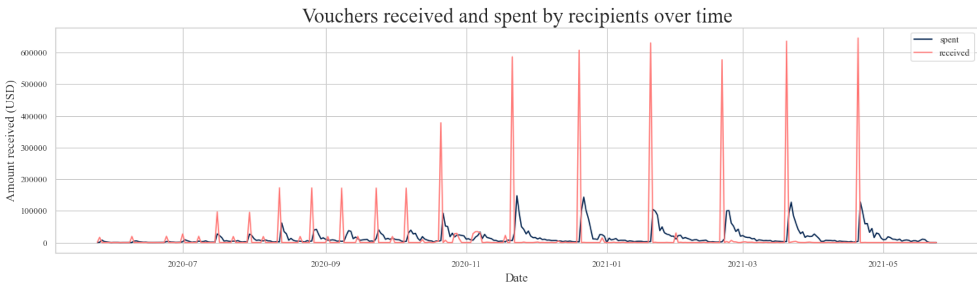


Figure 4.6: All vouchers received and spent by recipients over time

For a deeper inspection of the vouchers spent, two other variables on time have been investigated: the time of day and day in the week. For inspecting these variables in one glance, violinplots have been created on amount spent in time of day and day in the week. This violinplot has similarities with a box and whisker plot and shows the numerical data across categorical variables. The conclusion from the expenses at time of day in figure 4.7 is that there is a distinct difference amount spent on different days. Saturdays and Sundays have higher expenses in general compared to other days. Mondays and Fridays have remarkably high number of low expenses compared to other days. To show the amount spent on different times of the day, figure 4.8 displays a distinct difference in distribution of the amount spend from 8:00 am to 15:00 pm compared to 16:00 pm until 20:00 pm. By investigating these variables, we can assume that there could be different groups of customers shopping at various times of day and at different days in the week.

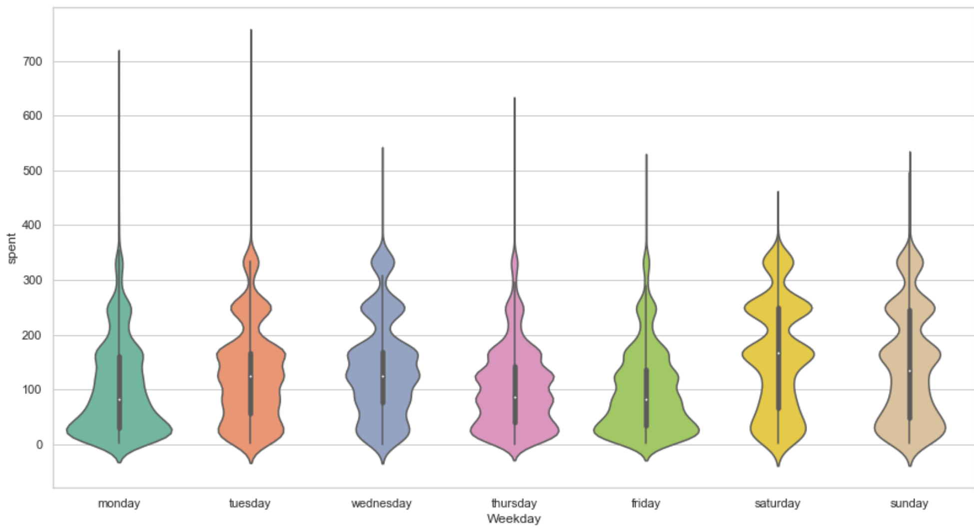


Figure 4.7: Voucher spent by recipients per day in week

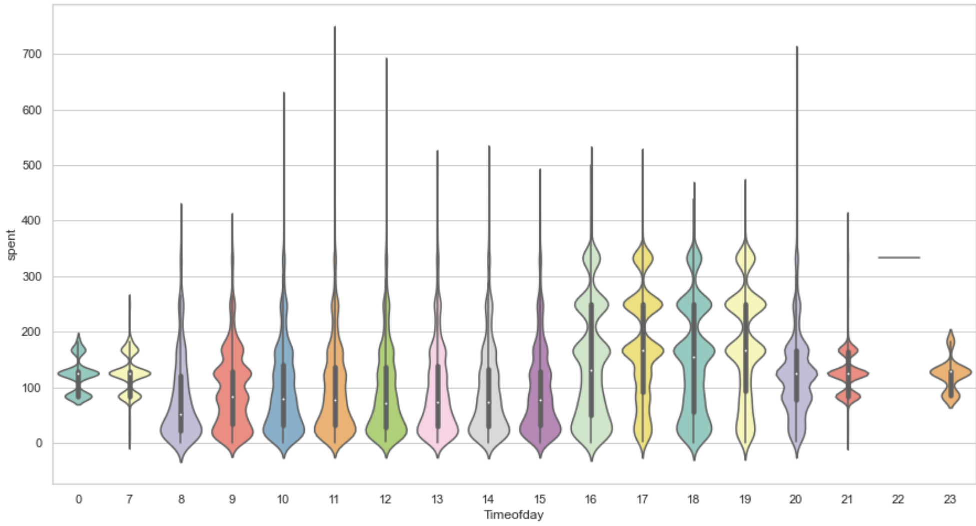


Figure 4.8: Voucher spent by recipients per time of day

5

Customer Segmentation

This chapter dives into customer segmentation theories to find common and useful methods which could be applicable on this research.

> The following subquestions are being answered in this chapter:
>
> - What customer segmentation methods are potentially applicable to this research?
> - What are variables from the field of customer segmentation that can be useful in describing behaviour of CVA recipients?
> - How can these methods be applied on the data of Sint Maarten?

5.1. Customer segmentation

Smith in 1956 was the first to formulate the term market segmentation, describing it as ”a way of an approach of looking at a heterogeneous market as a collection of smaller homogeneous markets in reaction to various product preferences among largest market divisions”. In light of ever-changing customer expectations and a fast changing globe, experts have delved further into customer segmentation as a one of the ways to segment markets (Croft, 1994; Lilien & Kotler, 1983; Weinstein, 1987).

Segmentation in social marketing, in particular, uses behavioural and psychographic data to help social marketers think about segment size and new ways to explore consumer characteristics. Social marketers can thus target large, sustainable, and accessible segments in the formulation of promotional and marketing activities. Marketers may use segmentation to carefully establish target customer groups and then focus available resources to the segments that are most promising (Rundle-Thiele et al., 2015). This revealing of customer behaviour for each segment is done in a variety of fields such as retail markets, tourism, e-commerce and transportation.

Previous research identified six primary indices to segment customers: demographic attributes, psychographic attributes, behavioural attributes, customer status, customer value, and RFM (Recency, Frequency, Monetary)(Aaker & Fournier, 1995; Kotler & Armstrong, 1999). Although the literature on customer segmentation is growing, there is not one best way to segmentate customers. “In part, there are literally millions of ways to divide up the market in any given context” (Aaker & Fournier, 1995).

5.1.1. Approaches

There are different approaches which can be utilized to do customer or market segmentation. These two approaches are most frequently used: common-sense segmentation and data-driven segmentation (Dolničar, 2004; Wedel & Kamakura, 2012). The biggest difference between the two approaches is the used variables. Common-sense segmentation, also known as a-priori segmentation, is a fairly simple way which uses one single variable to divide customers into groups (Myers & Tauber, 1977). Data-driven segmentation, also known as a posteriori or response-based, is a newer approach which uses an undefined number of variables to segment consumer markets. With the rise of big data in the marketing fields, the use of data-driven segmentation has increased enormously compared to common-sense segmentation

(Wedel & Kamakura, 2012). Data-driven segmentation is used in this research because of the data availability and the higher potential to explain segments with more variables. Often used variables in data-driven segmentation are related to demographics, psychographics, behaviour, RFM-model and Market Basket Analysis.

5.1.2. Segmentation variables

Despite the fact that many studies on market segmentation have been published, the variables used to segment the market can be combined in a variety of ways, and there is no broadly accepted combination of variables. The variables to include are largely dictated by the business or market undergoing analysis, and the outcome is highly impacted by the input variables (Frank et al., 1972). Because there is a big variety of variables which can be used, a distinction is based on observable and unobservable variables, and on general variables and product-specific variables. Observable in this context means that the variables can be measured directly and general points to the fact that variable is independent from a certain product or service (Jadczaková et al., 2013). Unobservable variables are generally collected via interviews and questionnaires. These distinctions with examples of variables can be seen in table 5.1.

Observable, general variables is the most widely used for segmentation and consists of demographic, geographic and socio-economic variables, frequently shortened to geo-demographics. Unobservable, general variables are often called psychographics and tries to capture the lifestyle, personalities and values of consumers. This is one of the hardest data to acquire, but is very profitable in commercial context when translating triggers into marketing actions. Specific observable variables are related to expenditures and consumption behaviour while specific unobservable variables are linked to perceptions of specific products or utility.

	General	**Product-specific**
Observable	Geographic, demographic, socio-economic, household size	Behavioural characteristics, usage frequency, loyalty
Unobservable	Personality, lifestyle, values, psychographics	Perceptions, elasticities, benefits

Table 5.1: Segmentation bases (Jadczaková et al., 2013; Wedel & Kamakura, 2012)

The data in this research consists of general and product-specific variables which both fall in the category observable. According to Jadczakova, these variables can make good description of different segments (Jadczaková et al., 2013). Thus, with the availability of geo-demographics and consumer behaviour, we will dive deeper into examples of these variables in next paragraphs.

5.1.3. Geo-demographics

There are many variables which fit in the category of geo-demographics. Geodemography combines demography (on human populations) with geography (spatial analysis) and sociology (aspects of subcultures). In short, geo-demographic variables are based on characteristics of people.

Age, gender and income are the most used demographic variables in consumer segmentation (Makgosa & Sangodoyin, 2018). Various age groups are found to having different buying behaviours, for example younger people care less about traditions and are more open to newer technologies in comparison to older consumers. Age can be an accurate variable to use in consumer segmentation (Bijmolt et al., 2004).

Gender is another frequent variable used in consumer segmentation studies. Men and women differ in many facets of their behaviour, including purchasing habits, information processing, interpretation and product choice (Cleveland et al., 2003). In the literature, gender is one of the most well-supported variables that causes variances (Yousaf & Huaibin, 2013).

Based on socioeconomic characteristics such as income and academic background, consumers are classified by economic conditions and social standing (Weinstein, 2006). Especially income is a descriptive variable when conducting consumer segmentation in many studies (Bijmolt et al., 2004). Reactions on price changes and decision-making processes differ between low and high income groups (Creusen, 2010).

5.1.4. Consumer behaviour

Variables in consumer behaviour are often divided into two distinct theories: the Recency, Frequency and Monetary analysis (RFM) and Market Basket analysis (MBA). Both theories and their possible application on this research have been explored.

Recency, Frequency and Monetary

Recency, Frequency, and Monetary analysis is a well-known marketing approach for analysing customer behaviour (Khajvand et al., 2011; Mulvenna et al., 1999). This is accomplished by looking at the expenditures of consumers using three criteria: (R) purchase recency, (F) purchase frequency, and (M) purchase amount in monetary terms.

Customers who spend more money or buy more frequently, according to theories and studies, can be segmented in highly descriptive groups (Hu & Yeh, 2014). Consumers who have recently purchased also respond better to marketing campaigns than customers who have not just purchased. RFM metrics can be used for creating effective customer segments(Mulvenna et al., 1999). Another implementation of RFM is to calculate customer lifetime value; a description of how much a customer is 'worth' to a certain organisation. RFM analysis is used in various sectors such as telecom, tourism, banking, mobile phone operators, health care, retail store and supermarkets (Abbasimehr & Shabani, 2019).

Recency refers to how long ago a customer has made a purchase or was engaged in an activity. Frequency specifies how often a customer makes transactions in a certain time period. Monetary value is used for indicating how much a customer has spent at a specific store, sometimes even at specific products in a store.

Different variations of the RFM model exists such as the RFD, RFE and RFMI models (Maghawry et al., 2021). Recency, Frequency, Duration (RFD) can be used for analysing behaviour of readership or viewership oriented companies. Recency Frequency Engagement (RFE) is used to include other metrics which relate to customer commitment such as visit duration. Recency, Frequency, Monetary, Interactions (RFMI) is used to control the effects of different interactions of organisations on customers.

The RFM analysis can be implemented via different methods: one can chose to divide customers into groups (regularly used by clustering in five groups) or one can chose to formulate fitting variables for each RFM category (Veynberg et al., 2018). Division by customers in groups are ranked from 1 to 5 on each category. Since there are 5 groups of each category, in total 125 groups exist in this implementation of the RFM model. These clusters are then used for further analysis with variables such as demographics and income variables. The other method consist of creating variables from consumer behaviour, also known as features, which can be brought into the cluster analysis directly together with the other variables (Miglautsch, 2000).

Various organisations created segments of customers with RFM which have a certain distinguishable characteristic (Abbasimehr & Shabani, 2019). Often used segments are for example: highly engaged customers who bought the most and most recent, loyal customers who visit often, highest paying customers who don't come often (so called whales), faithful customers are those who return often but do not spend a lot and first time buyers are customers who don't visit the store frequently (rookies). There are many more custom-made segments in examples of RFM applications which often point into the direction of high or low engagement, high or low value, high or low frequency and recent or old activity. In this research, RFM can be used to describe variables on spending behaviour.

Market basket analysis

Market basket analysis (MBA) is used for analysing products bought by consumers (Agrawal et al., 1993). Synonyms and related analyses include transactional analysis, receipt analysis and acquisition pattern analysis. This is helpful in discovering deeper insights in the behaviour of customer even when customers are not aware of the presence of unexpected behaviour patterns. Market basket analysis is often used to predict and recommend items to users of a product or service. With the help of association rule mining algorithms, common combinations of products are explored using data from receipts (Aguinis et al., 2013). For example, if a shopping basket frequently contains product X and Y, it is likely that new customers with X in their basket will also purchase product Y. Subsequently in market basket analysis, user-based collaborative filtering algorithms are used to find other customers with similar baskets. These algorithms are not exclusively used for analysing visitors of supermarkets since these have also found their way into recommendation systems for streaming services and fraud detection.

In this research, market basket analysis could be applied on identifying similar consumers of vouchers by analysing their receipts. A hypothesis can be constructed on the existence of a relationship between similar users their and baskets. If there is a significant relationship between market baskets compositions among groups of recipients, more research could be conducted on the possibility of early monitoring and predicting vulnerability of a certain recipient. It has been chosen to not investigate the use of MBA any further in this research, because customer segmentation is a necessary first step before applying MBA.

5.2. Application of customer segmentation

The application of the earlier mentioned data-driven segmentation can be done in four ways. It can be conducted *a priori* (when the segments are decided in advance) or *post hoc* (when the number and specifics of segments are derived from the analysis) (Jadczaková et al., 2013). Two other ways are to conduct segmentation with descriptive or predictive methods. Descriptive methods explain relationships in a set of variables while predictive tries to find and explain relationships between dependent and independent sets of variables. Post hoc descriptive segmentation can be done via clustering methods and dimension reduction techniques such as factor analysis. Post hoc predictive segmentation is often done via classification and regression models (Wedel & Kamakura, 2012).

A general approach of these applications is to first cluster in order to get insightful customer segments (Birant, 2011). With the found customer segments and with the use of test data, a prediction model can be constructed to find classification rules. Classification rules can accurately place new customers into existing segments. Marketeers take this step even further with the use of recommendation models. Since these models use association rule mining algorithms combined with market basket analysis, they can predict association rules to classify products purchased together in different segments. During these steps, so-called personas can be made. These are detailed descriptions of a customer segment and are used for numerous objectives, also in the humanitarian sector.

Multiple papers use clustering with applying the RFM model, for example by implementing the K-means and DBSCAN algorithm (Monalisa & Kurnia, 2019). Aggelis and Christodoulakis applied cluster models using RFM scores on bank transactions (2005). Another research explored a generic architecture for performing customer segmentation using purchase data of six different products (Lefait & Kechadi, 2010). One paper explored expenditures over a period of time using RFM via time series clustering (D. Chen et al., 2015). These papers show that customer segmentation using RFM model via clustering approaches can be beneficial in describing customers. Therefore, this research will use data-driven segmentation approach using clustering models combined with variables on recency, frequency and monetary to explain consumer behaviour.

5.3. Feature Engineering

It can be concluded from the theory of customer segmentation that a data-driven segmentation approach using clustering has a high probability of being successful for categorizing recipients of evouchers. Also the RFM model can be benefited when fitted on existing data. For this reason, feature engineering is introduced to transform available data to relevant variables related to the aspects from the RFM model. Since the RFM model does not give prescribed features we can use, we have to come with relevant features on recency, frequency and monetary elements. This section outlines the transformation while the Jupyter Notebooks go into detail on how these variables have been constructed. An overview of all the available variables can be found in table 7.1.

5.3.1. Recency

Variables on recency are based on how short ago a recipient has been active. In evoucher projects this could tell something on how fast a recipient acts when they have received an evoucher. For this reason, a variable on the time between the topup and first purchase have been constructed. Another variable on recency is the number of days the evoucher has been used. This gives an indication of how active a recipient was when compared to other recipients.

5.3.2. Frequency

Frequency relates to how many times something occurs in a certain time frame. Recipients who purchase often are naturally considered more frequent users of evouchers. From the exploratory data analysis in chapter 4 it has been found that there are distinct differences in purchases on different days and during different times of day. Therefore, variables on the most visited time of day and day in the week have been composed from the voucher data. Another variable on frequency is the most frequently visited supermarket for each recipient. The number of visits to the supermarket has also been created as a variable on frequency.

5.3.3. Monetary

The variables on monetary aspects relate to the amount spend. A minimum and maximum purchase amount for each recipient has been derived from the data. Also, the sum of all expenditures, the average expenditure and median of expenditures are considered monetary variables in this research. The supermarket with the highest sum of all purchases of a recipient is selected as a monetary variable.

5.3.4. Limitations

The creation of variables on RFM have two limitations. The first is that variables related to time have been translated to one value, while they could have been presenting a certain trend over time. This inevitably led to loss of relevant information and can influence the outcome, especially when considering variables using modes (most occurring value in a series). Secondly, these variables have been chosen using background knowledge on the case study of Sint Maarten. There are multiple other variables who could also have been chosen as relevant RFM variables. The consequence of this could be a lower degree of reproducibility on other cases since these variables could be only relevant to this case.

6
Modelling

This modelling chapter exists of two different methods of reducing the dimensions of the dataset. A factor-cluster approach is being presented.

The following subquestions are being answered in this chapter:

- What are important aspects to consider when selecting segmentation models?
- What models are there available and do these compare?
- How can the model be evaluated and validated?

6.1. Curse of Dimensionality

The reason why the data first needs to be reduced before clustering is because of the Curse of (high) Dimensionality. Although this phenomenon sounds like the sequel of a Pirates of the Caribbean movie, it has important implications for this research. Bellman first coined this problem when he referred to the exponential growth of hypervolume when more dimensions are added (Bellman, 1961). Especially clustering algorithms suffer from this curse of dimensionality for multiple reasons. First of all, computational cost increases exponentially with every extra dimension. Secondly and most importantly, data space becomes sparse and this will make clustering much more inaccurate.

This can be explained more elaborately when imagining an example of four datapoints in three different dimensions, with a scale of one to five with intervals of one (by means of example). In only one dimension, the five data points are presented on a line, which creates a total density of 5/8. When considering two dimensions, the five data points can be presented in a grid, with a total of 125 regions with a density of 5/125. In a three dimensional space, the five data points can be presented in a matrix of 625 regions with a total density of 5/625.

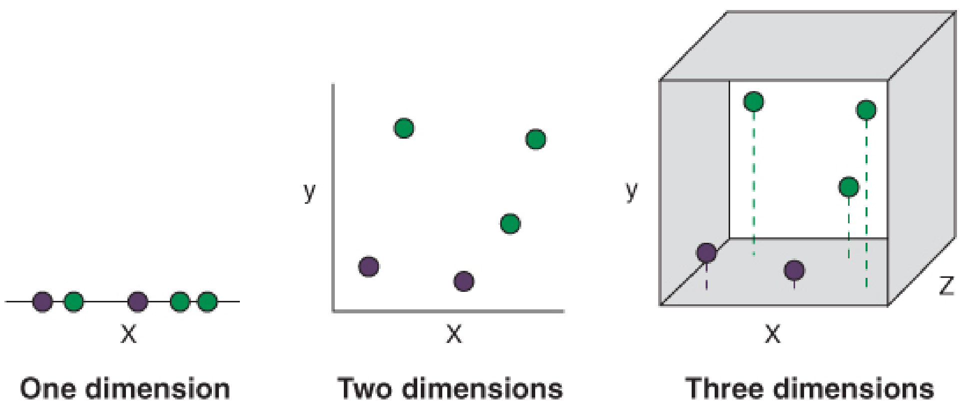


Figure 6.1: Curse of Dimensionality (Rhys, 2020)

This small example clearly shows that when dimensionality increases, the volume of the investigated area increases exponentially, but when data points do not increase correspondingly, this makes our coverage of the area of interest exponentially sparse. Conceptually this data will get lost in space. Since clustering algorithms (often) work with calculated distances between points, many dimensions make it look like each observation in the dataset is equally far away as can be seen in equation 6.1. The concept of distance becomes less precise because the distance between any pair of points converges. Clustering algorithms will conclude that each datapoint is its own cluster and this will not give meaningful results.

$$\lim_{d \to \infty} \frac{d_{max} - d_{min}}{d_{min}} = 0 \tag{6.1}$$

Machine learning methods are statistical by nature and most models just count number of observations in a certain space. ML models, and especially clustering models, can get lost in space with a small amount of data points with many dimensions. There are a few solutions for working with high dimensional data. One can choose to initiate a specific type of clustering used for high dimensional spaces such as projection-based clustering or projected clustering or one can try to reduce the dimensions before clustering.

6.2. Dimensionality Reduction

Reducing dimensions before clustering can be done via dimensionality reduction methods such as factor analysis and principal component analysis. The goal is to find underlying trends in order to combine variables into a new frequently called latent factor. When we have a lot of dimensions, the true dimensionality is often much lower since some variables can be explained in this latent factor. We will first try factor-cluster segmentation since this has been used in literature before (Ko et al., 2018; S. Smith, 1989). But factor-cluster segmentation has its downsides since a high explained variance is often removed and the interpretation of factors can be misleading (Dolnicar & Grun, 2011). Therefore, also other methods are considered for implementation in this research.

6.2.1. Considered Methods

Principal Component Analysis and Factor Analysis

Principal Component Analysis (PCA) can be used as dimension reduction technique too by capturing variance in a smaller set of variables. However, there are significant differences compared to factor analysis. The approach of PCA is to create one or more components by using a linear combination of input variables. PCA is built to optimally capture variance into one or more principal components. A conceptual example can be seen in figure 6.2.

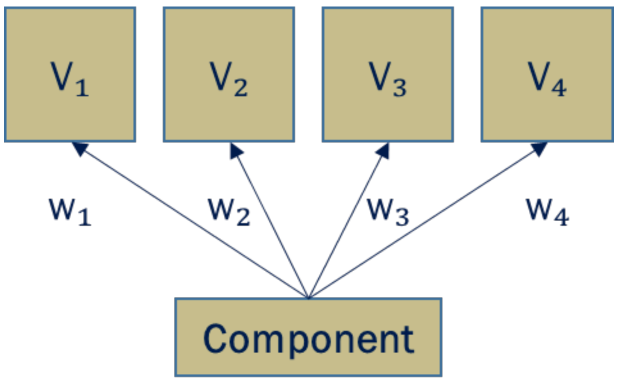


Figure 6.2: Conceptual workings of principal component analysis

This conceptual model can be described as a simple equation where w_i are the specific weights of each variable V_i. Usually more components C are being considered for capturing the most variance.

$$C = w_1V_1 + w_2V_2 + w_3V_3 + w_4V_4 \tag{6.2}$$

Factor analysis approaches dimension reduction in a different way. Where PCA optimizes the amount of variance captured in a set of components, FA assumes variance is explained by a latent factor. This underlying factor cannot be directly measured in a single variable but is explained by a set of variables.

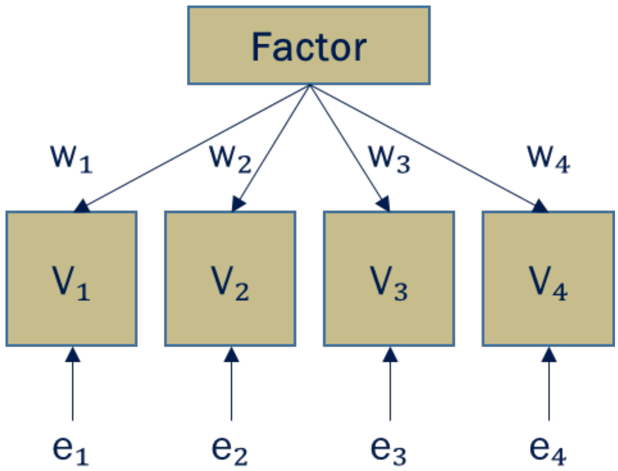


Figure 6.3: Conceptual workings of factor analysis

This conceptual model can be interpreted as a set of regression equations where w_i are the specific weights of each variable V_i. Since some variance will be unexplained by the factor, error terms e_i are assigned to each variable.

$$\begin{aligned} V_1 &= w_1F + e_1 \\ V_2 &= w_2F + e_2 \\ V_3 &= w_3F + e_3 \\ V_4 &= w_4F + e_4 \end{aligned} \tag{6.3}$$

For this research, we are able to use factor analysis or principal component analysis for dimension reduction. Although, since this research will utilize a cluster model, it would be preferred to use meaningful variables as input. Principal components are eventually just artificial variables capturing the most variance of a set of variables, so they don't have to be meaningful. Factors are often meaningful variables and are therefore preferred to use during cluster analysis for interpretation of results. Even though PCA usually conveys more information in the components compared to factors in FA, it has been chosen to seek for meaningful variables from exploratory factor analysis.

Other methods considered

More methods on dimension reduction exist and are briefly considered for applying in this research. For example, variations of PCA and FA which work with categorical variables are looked into. Examples of these methods include FAMD (factor analysis of mixed data) and PCAmix (principal component analysis of mixed data). Also methods of using only categorical data are examined such as MCA (multiple correspondence analysis for categorical data) and CATPCA (categorical PCA). For the sake of demarcation, only PCA and FA have been selected to consider as feasible dimension reduction methods for this research.

6.2.2. Factor Analysis

From the comparison between factor analysis and principal component analysis, it was found that factor analyses produces meaningful factors while principal components are dimensionless and often artificial. Therefore, factor analysis is the considered method for further dimension reduction. Factor analysis is a method to take correlations between variables away and replace collections of variables for a chosen number of factors (Williams et al., 2010). This is useful for when there is an assumption that the many correlations between variables can be described by a (collection of) latent factor(s). This technique is often used for combining questions from surveys to find an underlying factor which explains most of the answers of a selection of questions. This can be also useful for exploring data for patterns. We can theorize that certain variables move together for various reasons and factor analysis is a helpful method to investigating if this variable reduction is possible. There exists two kinds of FA: exploratory factor analysis and confirmatory factor analysis. The first type of factor analysis is used when the researcher has no expectation of the number of latent factors to be found (Hadi et al., 2016). Confirmatory factor analysis is used for verifying a certain structure in the data. Since this research aims to explore if certain factors can be constructed, exploratory factor analysis is appropriate to use.

Assumptions

The following assumptions are being made when conducting factor analysis:

- There are no outliers or missing data
- Sample size must be greater than factor
- Variables should be continues variables (interval and/or ratio)
- There are expectations of a latent factor describing the observed data
- Variables must be interrelated (Bartlett's Test of Sphericity)
- Correlations between variables are linear
- There is no multicollinearity

A recommended ratio of variables and observations is $5*2^k$ or at least 2^k (Formann, 1984). With our dataset of approximately 3,000 observations and around 70 features, we should strive to limit the number of features to 9 features and at least 11 features for running factor analysis.

Statistical tests, rotations and loadings

Before a factor analysis can be conducted, the data needs to pass some statistical tests and needs to check for correlations. A correlation matrix can be constructed to see if variables correlate. Without (linear) correlations between variables, factor analysis could not be considered appropriate. The Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) test are frequently used tests to check whether the data is suitable for factor analysis. The Bartlett's Test of Sphericity needs to test significant, pointing to a high probability of finding factors. The outcome of the KMO test ranges from 0 to 1 with a high number indicating good applicability for factor analysis with the used data. Kaiser himself suggested a bare minimum of 0.5 to test adequate on the KMO test (Kaiser, 1974). Values above 0.6 are considered sufficient in similar factor analyses used in marketing research (Pallant, 2020).

When the data passes the statistical tests, the factors can be extracted from the data and eigenvalues are computed. The eigenvalues of factors illustrate the amount of variance the factor explains. A greater eigenvalue than 1 means that a factor explains more variance than a single variable. A scree plot is used to visualize the eigenvalues per number of factors considered. Using this scree plot and the elbow method, we can determine the proper number of factors. When the number of factors have been chosen, the factors can be interpreted by using loadings and variance. The loading illustrate how much a factor explains a variable, where the loading score will range from -1 to 1. A loading score close to 0 demonstrates that factor does not have a significant influence on this specific variable. Extracting factor loadings can be optimized using different rotation methods. This research used the Varimax rotation to maximize the loadings on variables, resulting in a high loading score for a small number of variables. The amount of variance explained by choosing the number of factors is essential to get an understanding of the usefulness of the factor analysis. To reduce a high number of variables to a lower number of factors, a high percentage of total variance explained is desired to minimize the loss of explanatory power of variables.

Limitations

One of the limitations of using FA is that this only applies to numerical data. The categorical data in this research will not be considered in dimensionality reduction. Furthermore, interpretation of factors is subjective and should be done with background knowledge of the variables. Renaming factors to a meaningful factor can be hard and is sensitive for interpretation of the researcher. Another limitation is that deciding the number of factors is a debated field and there is not much agreement on one best method. Therefore, simulation and inspection of different outcomes of factor analysis can be enforced to create better understanding of the generated factors.

This research will apply factor analysis on the geo-demographic and behavioural data of recipients of evouchers. Data from the survey will not be used in factor analysis since this data will not be used as input for clustering. With factor analysis, the aim is to reduce as much of the variables to a smaller number of meaningful factors.

6.3. Clustering

Clustering aims to separate data into homogeneous groups. For example, clustering algorithms will combine two similar data points into a single cluster while separating dissimilar observations into other distinct groups. Unsupervised machine learning clustering algorithms are widely researched for continuous or categorical data (Preud'Homme et al., 2021). But this task gets more difficult when dealing with mixed data (i.e., data from nominal, ordinal, interval and ratio scale). The difficulty of performing mathematical operations concurrently to all types of feature variables is a major issue when clustering data containing mixed variables. The issue is determining the optimum distance or model to analyse both sorts of information simultaneously.

6.3.1. Considered methods

Luckily, several clustering techniques have been developed specifically for clustering categorical and numerical data. Ahmad and Khan have published a complete taxonomy of available approaches for clustering mixed-variable data sets (2019). In both simulated and real-world settings, Preud and colleagues evaluated the effectiveness of clustering techniques for mixed data (2021). They observed that model-based techniques and K-prototypes performed better with mixed data in most of the cases they looked at. Heterogenous data, also called mixed data, consist of both numerical and categorical data. This research includes categorical data (e.g., gender, district, living situation) and numerical data (e.g., income, age, vouchers spent). The K-prototypes clustering method and the Kamila algorithm are both extensions of the well-known k-means algorithm.

Another option on clustering mixed data is to use one-hot encoding categorical variables. One-hot encoding is used to convert categorical data to numerical data by two steps. First, unique values of a categorical variable are labelled with integer values. Secondly, every unique combination of categorical variables is considered as a new binary variable, often called dummy variable. For example: considering the categorical variable district has three different values 'Philipsburg', 'Simpson Bay' and 'Cole Bay'. This variable can be one-hot encoded into three different binary variables where only recipients from 'Philipsburg' are assigned 1 in the binary variable of 'Philipsburg'. This way of transforming categorical variables into multiple binary variables has been considered, but has been found inconvenient in this research. The reason for this is that with using new one-hot encoded variables, the number of variables increases immensely which is unfavourable for avoiding negative effects from the curse of dimensionality.

Hierarchical clustering has also been considered for clustering categorical data. Using the Gower distance, a distance matrix can be calculated working with the categorical variables (Akay & Yüksel, 2018). Hierarchical clustering can be done both divisive as agglomerative, using a top-down and bottom-up approach respectively. With the help of a dendrogram, different cluster configurations can be discovered. The downside of this method is that this can be very slow and is very sensitive when data contains a high level of error. Therefore, hierarchical clustering is not further considered.

This research uses k-prototype because of the high degree of applicability of a non-hierarchical clustering algorithm which has been designed for numerical and categorical variables.

6.3.2. K-prototype

For clustering mixed-type data, Huang developed the k-prototypes method, which combines the principles of the k-means algorithm and the k-modes algorithm (Huang, 1997). K-means is used on numeric data only and works with finding minimal Euclidian distances between points (Iam-On et al., 2014). K-modes has been developed for categorical data on finding the most common (mode) on a set of categorical variables to cluster into similar segments (MacQueen et al., 1967).

Theory

The following formula gives the dissimilarity between two observations x and y, for $d1$ numerical features and $d2$ categorical features:

$$D(x,y) = \sum_{j=1}^{d1}(x_j - y_j)^2 + \gamma_1 \sum_{j=d1+1}^{d2} \delta(x_j, y_j) \text{ where } \delta(x_j,y_j) = \begin{cases} 1, & if\ x_j = y_j \\ 0, & if\ x_j \neq y_j \end{cases} \tag{6.4}$$

The first part of this equation stems from the K-means algorithm which calculates the square of the Euclidian distance. The second part adopts the Hamming distance which is defined by the δ

simple matching distance (Hamming, 1950). This combination of k-means and k-modes creates a new hybrid cluster centre called the prototype. An addition of γ_1 is used for balancing the influence of the categorical variables to numerical variables.

The goal of this algorithm is to minimize the following Cost-function:

$$P(U, Z) = \sum_{i=1}^{n} \sum_{l=1}^{k} u_{il} D(x_j, z_l) \tag{6.5}$$

To reduce the value of the Cost Function, the k-prototypes algorithm separates the dataset into distinct subclusters. The process behind the K-prototype algorithm can be described with the following steps:

1. A randomly selected k initial cluster centres are selected from the dataset.
2. With equation 6.4 the dissimilarity is calculated between x_j and y_j. This data point will be allocated to the nearest cluster.
3. Using the current cluster centres, the dissimilarity of the data points are being recalculated. The data points are reassigned to the nearest cluster using the method of average value for numerical data and most frequent value for categorical data. The cluster centres are updated after this.
4. Steps 2 and 3 are repeated until the cost-function is stable and no longer declining.

This algorithm can be used for simulating the total cost of k number of clusters. Visualizing the total cost of k clusters in a scree plot is helpful to find an optimum number of clusters which no longer account for a significant change in total cost. This should be done with care, since the found clusters should represent a meaningful segment of the data points. Finding an appropriate number of clusters is therefore not only a mathematical exercise, but also a subjective examination of finding relevant clusters. Without the known number of clusters beforehand, we have to simulate the cluster algorithm over a large number of potential clusters. No definitive method choosing the number of clusters is considered superior (Hennig & Lin, 2015; Tibshirani & Walther, 2005). The optimum number is chosen rather arbitrarily and based on the scree plot and by using the elbow method. An additional check of choosing the number of clusters can conducted by using validation indices.

Although k-prototypes is considered as one of the more effective ways to cluster with mixed datatypes, it also has its limitations. One of this is that the simple Hamming distance can result in loss of information and thereby not projecting the real situation between cluster centres and data points. Another limitation is that parameter γ needs to be manually determined and that the initial clusters are randomly selected. This creates some uncertainty and randomness in cluster determination. One way to deal with these shortcomings is to simulate with different parameters and compare outcomes.

Another common limitation with clustering is the algorithm being trapped in a local optimal solution while there is a better global solution available. The initial cluster prototypes have influence on the probability of being trapped in a suboptimal cluster. There are some methods on solving this problem, for example by introducing deviations of the initialization phase to come up with new initial cluster centroids. But still then, it cannot be guaranteed that the cluster algorithm will be the global optimal solution. However, there is a technique for solving this with genetic algorithms and simulated annealing, but this is out of scope in this research (Goldberg, 1989; Kirkpatrick et al., 1983).

Implementation

When implementing k-prototype, the numerical variables should be scaled to the [0,1] interval. This is being done by a min-max features scaling equation where x is the original value and x' is the scaled value. Scaled values make sure every feature is of the same importance in the clustering algorithm.

$$x' = \frac{x - min(x)}{max(x) - min(x)} \tag{6.6}$$

The implementation of k-prototypes repeatedly calculates for every k number of clusters the total cost. The 'Huang' initialization function of findings centroids refers to the selection of most frequent observations as first centroids (Huang, 1997). To find a right fit of γ for balancing numerical and categorical variables, a simulation of γ has been carried out with values from 0.5 to 2. Huang indicates that a suitable γ usually lies between 1 and 2 but that more research should be done to find strong rules to construct γ (1997). The K-Prototype algorithm in R calculates γ so that the numerical and categorical variables are equally weighted. This standard build in calculation of γ in the used software package is used for this specific case. Implementing k-prototypes can be done via the ClustMixType package in R (Vos, 2021). Python can be used for analysing clusters with visualizations.

K-prototypes Evaluation

Evaluating cluster performance can be done by different validation indices and by expert validation. Evaluation is important because algorithms are blind and can lead to ungrounded answers. Expert validation is done by someone in the field with a lot of background knowledge on the input data. This kind of validation helps to argue the validity of the found clusters by inspecting by hand and deciding for each cluster the likelihood of being a valid cluster.

For mixed datatypes the Dunn Index, McClain Index and Silhouette Index are most useful (Aschenbruck & Szepannek, 2020; Chouikhi et al., 2015). In theory, another simple evaluation can be done by visualising clusters in a 2D or 3D plot by using one of the earlier mentioned dimensionality reduction method called PCA. However, interpretation will be hard when visualizing this result onto dimensionless and artificial components.

The ClustMixType package in R has a build-in implementation of different validation indices. The Silhouette, Dunn, Gamma and G+ index are available to directly use on the cluster results. Other indices need a separate custom-built script in order to implement this on K-prototypes. Since, Dunn, Gamma and the G+ index are too computationally expensive, it has been chosen to only implement the Silhouette index in this research for validating results with indices.

The Silhouette Index works with a Silhouette Coefficient which is a measure of how close each point in a cluster is to neighbouring clusters. This gives insight in the separation distance between clusters. For a single data point, the Silhouette Coefficient s is the following:

$$s = \frac{b - a}{max(a, b)} \tag{6.7}$$

Where a is the mean distance between the specific data point and all the other points in the cluster and b is the mean distance to all other points in the closest neighbouring cluster. The Silhouette index results in a score on a scale from -1 to 1, where we want to maximize the score (Foss et al., 2016). When calculating the Silhouette Coefficient for all data points of k clusters, an average Silhouette Score can be calculated. The average silhouette coefficient k is determined by the following equation:

$$k = \frac{1}{n}\sum_{i=1}^{n} \frac{b_i - a_i}{max(a_i, b_i)} \tag{6.8}$$

7

Results

This chapter presents the found factors, clusters and their evaluation indices. This chapter starts of with a presentation of the available variables for clustering. Next, the factor-cluster approach leads to a specific set of clusters with their own distinct characteristics. These clusters are compared to the survey data to see if conclusions can be made on needs of recipients.

The following subquestions are being answered in this chapter:

- Which factors can be found using factor analysis?
- Which cluster characteristics are most distinguishable?
- What insights can survey data reveal when using the cluster results?

7.1. Variables

Table 7.1 gives an overview of all the variables with their original names, description and category which can be used for factor analysis and clustering. Due to the curse of high dimensionality, using all the variables will create unmeaningful clusters. Therefore, a subset of these variables has been chosen for the factor-cluster approach. The motivation for selecting the chosen subset is based on the most describing variables according to literature on customer segmentation. More combinations of subsets are possible and could be implemented in future research to see if big differences emerge in clusters. For now, this specific subset is further explored.

Demographic variables on age, gender and income are selected because of their frequent presence in customer segmentation studies and their recognized relevance. Also, one of the knowledge gaps on recipient characteristics motivated to further research the impact on female-headed households, so this is another reason to include gender in the subset of variables (CARE International & ideas42, 2020) One geographic variable on the district of residency has been selected in the subset. A seemingly other interesting geographic variable could be distance from home to the visited supermarket. But since Sint Maarten is a very small island and we don't have granular data, this is not a suitable variable for this specific case.

With feature engineering, new variables on recency, frequency and monetary aspects have been created during the step of feature engineering. A selection of these variables is included in the subset of interest. Monetary variables is the sum of all vouchers received and the sum of all vouchers spent. Another two monetary variables which are useful for segmentating are the maximum amount of money spent during a visit and the minimum amount spent. Lastly, a monetary variable on the difference between money spent and received has been selected because this could be predictive in how recipients use or don't use vouchers. The variable on frequency includes the number of times recipients went to visit the supermarket and also the most frequently visited supermarket. Since exploratory data analysis showed differences between time of day and day in the week of supermarket visits, these variables too are selected in the subset as variables on frequency. Recency variables include the number of days between first and last expenditure, since this could be explanatory in the problems recipients perceived.

Also, the time between first voucher received and first voucher spent could be an informative variable on the behaviour of recipients.

Since knowledge on the influence of household compositions is lacking, four different variables on household compositions are included in the subset. Literature points to extra vulnerable groups (e.g., female-headed households or medically disabled persons) to be included in research on CVA (CARE International & ideas42, 2020). The data on Sint Maarten includes the number of medically disabled persons and number of pregnant women in households. Since the number of medically disabled persons is less than thirty, it has been decided to not include this variable into the cluster analysis. Other variables that are not included in the subset were considered to have too many possible outcomes for clustering (for categorical variables), thus being less suitable for clustering.

Variable name	Variable description	Category	In subset
address_district	District (9 possibilities)	Geographic	Yes
age	Age of main beneficiary	Demographic	Yes
composition_pregnancy_amount	Number of pregnant women in household	Demographic	Yes
composition02	Number of people in household with age between 0 and 2	Demographic	Yes
composition1864	Number of people in household with age between 18 and 64	Demographic	Yes
composition317	Number of people in household with age between 3 and 17	Demographic	Yes
composition65	Number of people in household with age above 65	Demographic	Yes
days_used	Number of days between first and last expenditure	Recency	Yes
difference_topup_spent	Different between money received and spent	Monetary	Yes
finance_rent	Living situation (paying rent, homeless, without rent/mortgage, own without mortgage, own with mortgage)	Demographic	Yes
finance_total_income_USD	Total income per month (USD)	Demographic	Yes
gender	Gender of main beneficiary	Demographic	Yes
median_spent	Median of all expenditures of recipient (USD)	Monetary	Yes
mode_dayinweek	Most frequently visited day in the week	Frequency	Yes
mode_timeofday	Most frequently visited time on the day	Frequency	Yes
most_freq_supermarket	Most frequently visited supermarket	Frequency	Yes
spendings_count	Number of times in	Frequency	Yes
spendings_max	Maximum expenditure	Monetary	Yes
spendings_min	Minimum expenditure	Monetary	Yes
spendings_sum	Sum of all expenditures	Monetary	Yes
time_between_first_topup_and_transaction	Time between first received topup and first transaction	Recency	Yes
topup_count	Number of all expenditures	Frequency	Yes
topup_sum	Sum of all topups (money received)	Monetary	Yes
average_spending	Average expenditure (USD)	Monetary	No
composition_total_h_hmembers	Number of people in household	Demographic	No
finance_a_o_v_amount_USD	Pension per month (USD)	Demographic	No
finance_onderstand_amount_USD	Social assistance from government per month (USD)	Demographic	No
finance_s_s_r_p_amount_USD	Unemployment support per month (USD)	Demographic	No

finance_salary _amount_USD	Salary per month (USD)	Demographic	No
highest_sum _supermarket	Supermarket with highest expenses of recipient	Monetary	No
language	Language	Demographic	No
medical_disabled_amount	Number of medically disabled people in household	Demographic	No
modality_preference	Preference of aid modality	Demographic	No
topupamountmonthly_USD	Monthly amount topup received (USD)	Demographic	No

Table 7.1: All variables which can be used for clustering

7.2. Factor Analysis

The first step in the two-step factor-cluster approach is to reduce dimensions on the variables with the use of factor analysis. To find potential variables which can be combined in a factor, a simulation has been run over 100,000 times to randomly generate different subsets of the available data. The choice for 100,000 simulations was made to balance output usefulness with the high computational costs that come with many simulation runs. A condition for the creation of a subset was that the Bartlett's Test of Sphericity was found significant and Kaiser-Meyer-Olkin (KMO) test scored higher than 0.6. This ensured that the subsets found were appropriate for conducting factor analysis with. With the simulations, the most frequently occurring variables with a high score on the KMO test were selected as most promising variables for conducting a factor analysis on. Parallel to this simulation, a heatmap of all correlations between all available variables was generated which can be seen in figure B.1 in appendix B.

From the heatmap and the simulation of subsets a set of variables was found which has a high total explained variance when used in factor analysis. There exists a high (positive or negative) correlation between those variables, indicating a high chance for a meaningful factor analysis. Therefore, we chose to use this set of variables as highest input for our variables. Other variables did not show a significant correlation (in this case lower than 0.5) and are therefore not considered in the factor analysis. The following list of variables is used in factor analysis:

- 'composition_total_h_hmembers'
- 'topupamountmonthly_USD',
- 'finance_total_income_USD'
- 'finance_total_accumulated'
- 'average_spending'
- 'median_spent'
- 'topup_count'
- 'topup_sum'
- 'days_used'

With this subset of variables, the factor analysis was conducted to ultimately reduce the number of dimensions in the dataset. The Bartlett's Test of Sphericity was found significant so there is a high probability of finding factors. The Kayser-Meyer-Olkin tests is 0.63 which means this data is suitable for carrying out factor analysis. Based on the scree plot in figure 7.1, an optimum number of 4 factors has been selected.

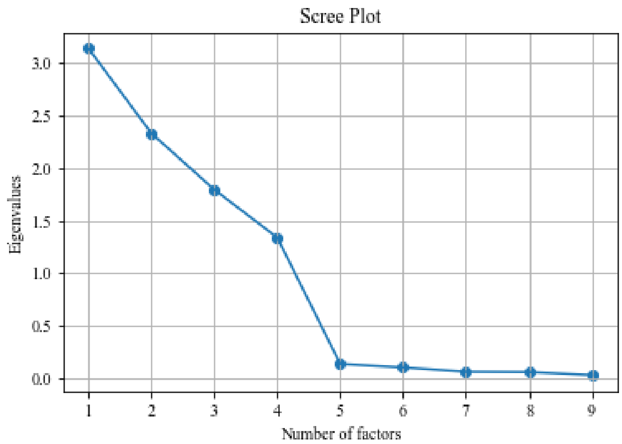


Figure 7.1: Screeplot of Factor analysis

The Varimax rotation is used to project loadings on the 4 factors. This Varimax rotation maximizes the variance shared among variables. The loadings per factor, using 4 factors, can be seen in figure 7.2. The scores higher than 0.9 and 0.8 have been highlighted to clarify the highest loading per factor, which can be explained as the characteristics of this specific factor. In total, these loadings explain 91.9 percent of the variance of this specific subset.

	0	1	2	3
composition_total_h_hmembers	0.089922	0.900173	0.132473	0.089350
topupamountmonthly_USD	0.101831	0.969165	0.113317	0.077501
finance_total_income_USD	0.024780	0.082655	0.026455	0.993412
finance_total_accumulated	0.008260	0.070415	0.009673	0.860733
average_spending	-0.002218	0.165873	0.983278	0.026521
median_spent	-0.034960	0.078900	0.936502	0.011386
topup_count	0.976330	0.005346	-0.029273	0.008958
topup_sum	0.852808	0.432310	0.066675	0.038683
days_used	0.988183	-0.010391	-0.060214	0.007458

Table 7.2: Factor loadings on each variable

It is useful to rename the factors to something more meaningful, so that cluster analysis can be interpretable with these factors. The following names were given to the newly found factors:

1. Factor 0: 'Long-term participation'
2. Factor 1: 'Big families and big receivers'
3. Factor 2: 'Average spent'
4. Factor 3: 'Total income'

The four factors were added to the existing dataset and the original variables were removed from this dataset. Implementation of these factors has been realized by min-max scaling the original subset of variables between 0 and 1, followed by multiplication of the factor loadings on the transformed subset variables. The sum of this product of factor loadings and min-max scaled variables, is considered the score on the factor for every recipient. This sumproduct of every factor score is again transformed to a scale from 0 to 1 for easy comparison between factors. For example, when a household scores high (close to 1) on factor 'Total income', this household has a high income, relative to other households.

The result of this factor analysis is a reduction of 9 variables into 4 factors. Although, this is not a high reduction in dimensions, this will help overcome the curse of dimensionality because a significant variance of 9 variables is captured in 4 factors.

7.3. Cluster analysis

Using the four new factors from the factor analysis, a variable selection for the clustering has been made. The K-Prototype algorithm constructs a certain number of clusters which can be analysed using different visualizations. The different clusters can be used for constructing personas as a representation of a certain group through common characteristics.

7.3.1. Selected data and cluster selection

Using knowledge from the case study and from customer segmentation theories, a subset of the variables and the 4 newly created factors has been assumed to explain a promising degree of explanation while clustering:

- 'address_district' (geographic)
- 'age' (demographic)
- 'gender' (demographic)
- 'finance_rent' (demographic)
- 'composition02' (demographic)
- 'composition317' (demographic)
- 'composition1864' (demographic)
- 'composition65' (demographic)
- 'composition_pregnancy_amount' (demographic)
- 'Total income' (monetary)
- 'Big families and big receivers' (demographic)
- 'Average spent' (monetary)
- 'spendings_sum' (monetary)
- 'spendings_max' (monetary)
- 'spendings_min' (monetary)
- 'difference_topup_spent' (monetary)
- 'Long-term participation' (frequency)
- 'most_freq_supermarket' (frequency)
- 'spendings_count' (frequency)
- 'mode_timeofday' (frequency)
- 'mode_dayinweek' (frequency)
- 'time_between_first_topup_and_transaction' (recency)

This subset is used in a simulation of running 2 to 30 clusters in the K-Prototype algorithm. For each number of clusters, the total cost is calculated and plotted on a scree plot, as explained in section 6. From the scree plot using the elbow method, the optimal number of clusters can be determined, which is 5 for this specific subset, because more clusters will not add a significant amount of extra information. The 5 different clusters are assigned to the original data and an analysis is conducted to see the characteristics of each cluster.

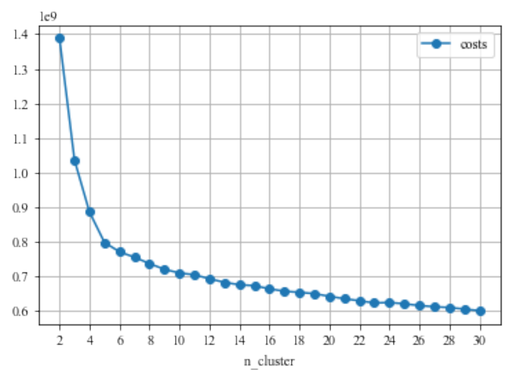


Figure 7.2: Scree Plot of Clustering Simulation

Internal validation of the found clusters can be performed using the Silhouette Index. This gives an indication of how close each point in a cluster is to neighbouring clusters. The averaged Silhouette Index for the simulation of 2 to 30 clusters is shown in figure 7.3. With 5 clusters, the Silhouette Index is around 0.48 which indicates that on average, data points are well matched to its own cluster and badly matched to neighbouring clusters. This cluster result can therefore be called relatively compact. External validation was done by inspection by an expert from the field. After careful investigation of the generated clusters and their probability of being plausible, it was determined that the cluster analysis very likely represents existing groups and even some new groups during the project on Sint Maarten.

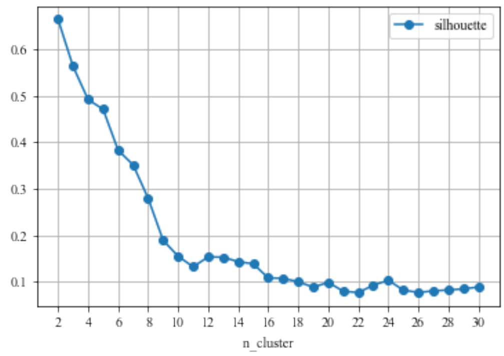


Figure 7.3: Silhouette Index of different cluster results

7.3.2. Analysis of clusters

The analysis of the found clusters can be done in different ways. First, the characteristics of each cluster relative to other clusters are inspected. This is done by using the mean of each cluster on each numerical variable and the mode for each categorical variable. The numerical mean values are then transformed to z-scores to be able to see outliers, and therefore certain unique cluster characteristics. These z-scores are plotted in a so-called radar plot, also known as spider plot. Figure 7.4 gives an overview of all the plotted z-scores on each variable, while table 7.3 present the number of recipients per cluster.

Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Number of recipients	735	360	807	332	637
Percentage	25.6%	12.5%	28.1%	11.6%	22.2%

Table 7.3: Observations per cluster

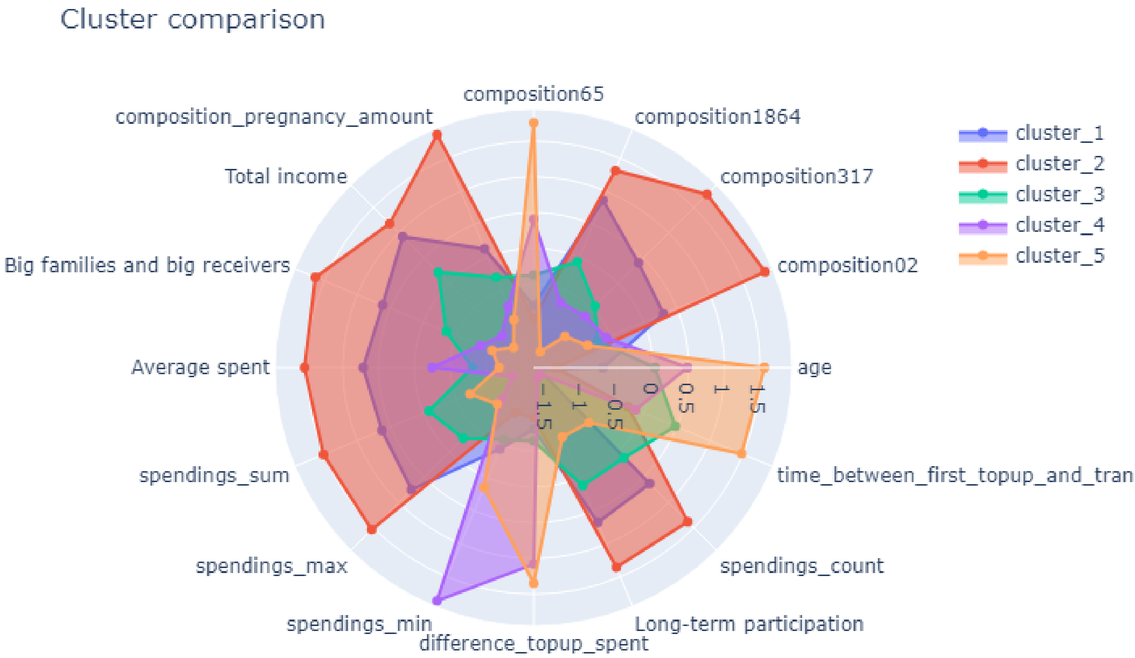


Figure 7.4: Radarplot of all clusters using z-scores of means

More cluster analysis has been included in appendix C. The distribution of categorical variables for each cluster can be seen in this appendix. Also figure C.7 until C.11 shows radarplots of means using absolute values instead of z-scores. The cluster analysis has made use of these figures in appendix C but most of these graphs have not been included in this section. This sparingly use of graphs was chosen to make this section as readable as possible and to only highlight the most distinctive characteristics, while not showing the least distinctive variables.

Next to the cluster analysis on variables used in the clustering algorithm, also variables not used in K-Prototype are analysed. These variables did not influence the algorithm in any way, but it can be still interesting to see if these variables differ among clusters from the average or mode of this specific variable.

Cluster 1: Families with relatively long participation in the voucher program

Cluster 1 consists of households with a relatively low number of elderly people. These households participate for a long time in the program and they react fast when receiving a new voucher. The main beneficiary often has a relatively young age. The other variables of this cluster are among the average of other participants in the voucher project.

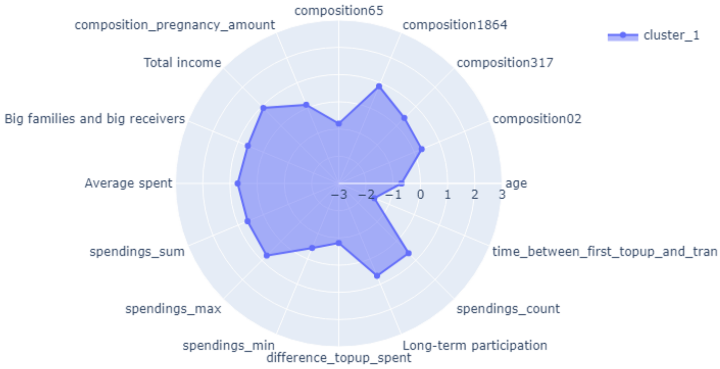


Figure 7.5: Radarplot of cluster 1

Recipients from cluster 1 had relatively few English speaking recipients, since 54% speak English compared to 58% of the population. Although the difference between other clusters who spoke was not big, this is still a noteworthy deviation from other clusters who have the same distribution in languages spoken. Also, since this variable is not used in the clustering algorithm, the difference among clusters is even more exceptional. Less proficiency in English in cluster 1 could lead to different communication problems with the staff of the voucher project, however this can not be said with certainty and should be further researched using the survey data. Another distinction in cluster 1 using variables who were not an input for the clustering model, is about medically disabled people. There were not many medically disabled people in the Sint Maarten Project (13 in total), of these people none were in cluster 1 and 2. The sample size is too small to make strong conclusions, but this could point to cluster 1 and 2 households having a lower probability of consisting of medically disabled persons.

Cluster 2: Young families actively using evouchers

Cluster 2 is characterized by household with many members, especially children. The main beneficiary is most likely to be among the youngest of the main beneficiaries. This group pays rent more often than other groups and thus own relatively less houses which is noteworthy, given the fact that they belong to a relatively high income group. Their behaviour in the evoucher project is best described by very active since they often go to the supermarket compared to other clustes and they have a small difference in the amount of vouchers they received and on the total expenses. Their shopping bags are among the most expensive and at other times they could be among the least expensive, compared to other clusters. Remarkably, this group has most visits to the supermarket on Thursdays while having the least on Sunday, when comparing this to all the other clusters. Also, this group tends to avoid the Fairway supermarket compared to other supermarkets and clusters. These young families have the highest likelihood of having newborns next year.

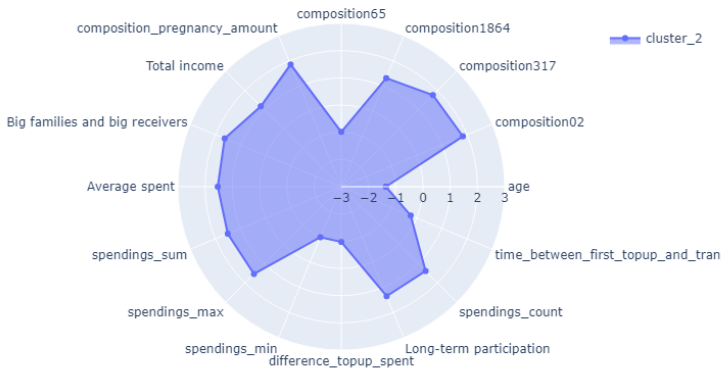


Figure 7.6: Radarplot of cluster 2

A distinctive characteristic of this cluster, when looking at variables not being used as input for the clustering algorithm, is the preference of aid modality. In cluster 2, 69% of the recipients preferred evouchers compared to an average of 92% for the population, which can be seen in figure 7.7. One of the explanations for this is the relatively high number of people with unknown preferences (value equal to 'missing'). Although around 28% of the people from cluster 2 have an unknown preference, this is extremely high compared to an average unknown preference of 5%. Preference is not used in the clustering model, but there still happens to be a big difference in cluster 2 which is peculiar. A reason for this could not have been found during this analysis, but could be further researched when applying correlation analyses on both numerical and categorical variables, while looking for relationships with preferences.

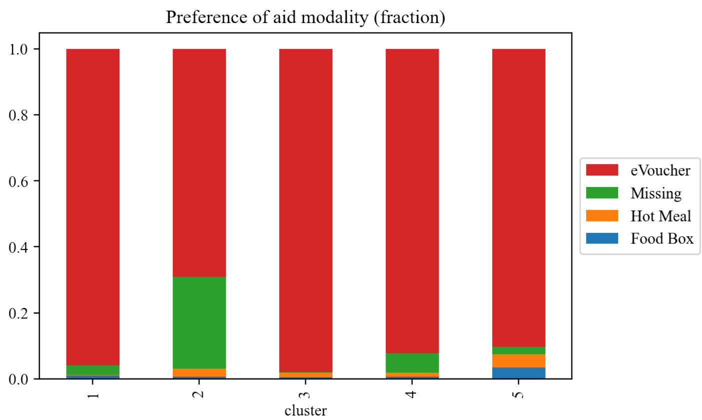


Figure 7.7: Preference of modality

Cluster 3: Average users

Cluster 3 is a typical average user. They don't have a specific characteristic compared to other clusters. Although, this is the biggest group with 807 members, they are their own cluster on being an average household participating in the evoucher project. The appearance of a cluster with no distinct characteristics can emerge during cluster analysis. There is a possibility that more subclusters exist inside this cluster, but they are not as distinct as cluster 1, 2, 4 and 5. This could be further substantiated with indices on the degree of separation, compactness and cohesion inside this cluster.

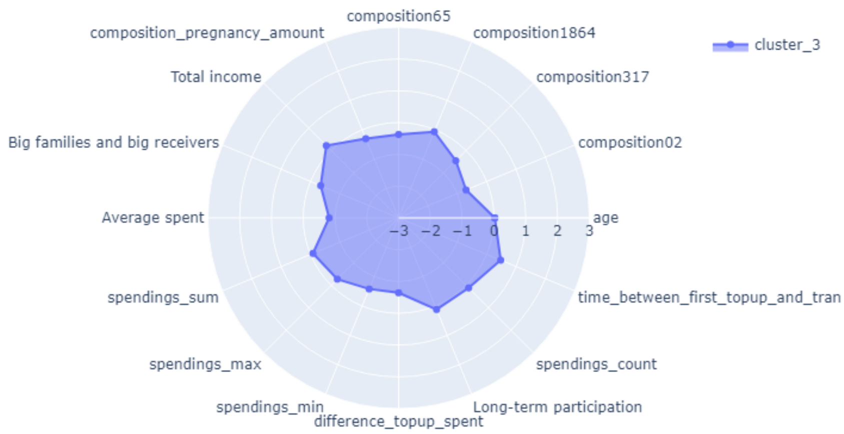


Figure 7.8: Radarplot of cluster 3

Cluster 3 is interestingly enough also a cluster in which the recipients district of residence is very comparable to the average distribution of the population, which can be seen in the row of cluster 3 and of the 'Average' percentage in table 7.4. Other clusters have outliers in their proportion of people living in certain districts. For example, cluster 4 has relatively many people living in Philipsburg and cluster 2 has relatively a higher precentage of the people living in Little Bay compared to other clusters. Cluster 3 is also regarding districts of residence quite close to average and not distinctive.

address_district	Cole Bay	Cul de Sac	Little Bay	Lower Prince's Quarter	Philipsburg	Simpson Bay	Upper Prince's Quarter
cluster							
1	18.20	19.60	10.60	30.10	5.20	2.60	13.60
2	18.60	16.40	13.90	30.80	6.70	2.50	11.10
3	17.00	20.40	9.40	29.50	6.40	3.20	13.80
4	16.00	17.20	6.60	29.80	17.80	3.00	9.00
5	18.50	21.70	8.30	30.50	6.10	3.50	11.30
Average	17.66	19.06	9.76	30.14	8.44	2.96	11.76

Table 7.4: Percentage of recipients in district per cluster

Cluster 4: Short-term users

Cluster 4 consists of slightly older than average people, with a lower income and with a short term participation in the project. The times they used the vouchers, it was mostly used in high amounts compared to other users. Notably, people from cluster 4 have a relatively high chance of living in district Philipsburg compared to other clusters.

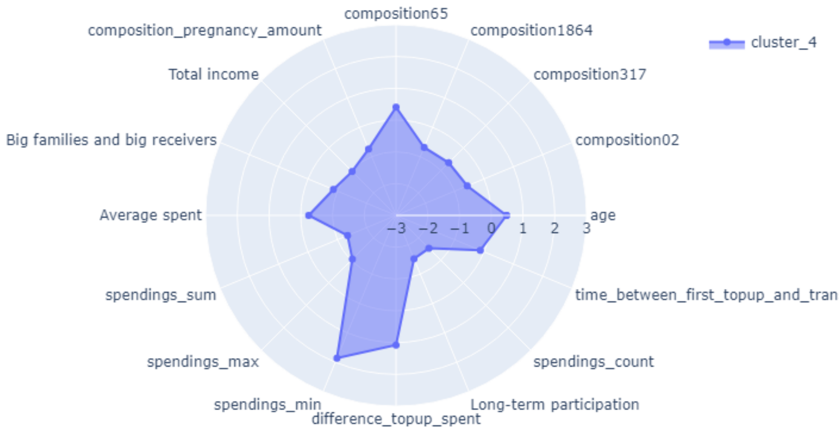


Figure 7.9: Radarplot of cluster 4

During the factor analysis, the variable on income was used, among others, to create a new factor called 'Total income'. During this process, the original variable on income was removed, but we can still analyse the original variable in different clusters. One thing that stands out is that 47% in cluster 4 did not earn any salary compared to 25% in cluster 1, 29% in cluster 2, 29% in cluster 3 and 37% in cluster 5. This is a relatively high number of people not earning any salary and this is an indication of a high ratio of recipients in cluster 4 being unemployed. Further analysis could point into how unemployed people experienced using these vouchers.

Cluster 5: Older people living alone
Cluster 5 are the older couples or older people living alone. They had a relatively long waiting time before they spend their first voucher in the supermarket. People from cluster 5 did not spend all their vouchers, which is interesting given their relatively low income, but could be explained by their short participation in the program. There are not many homeless people using vouchers in this case study, but cluster 5 has the highest number of homeless people.

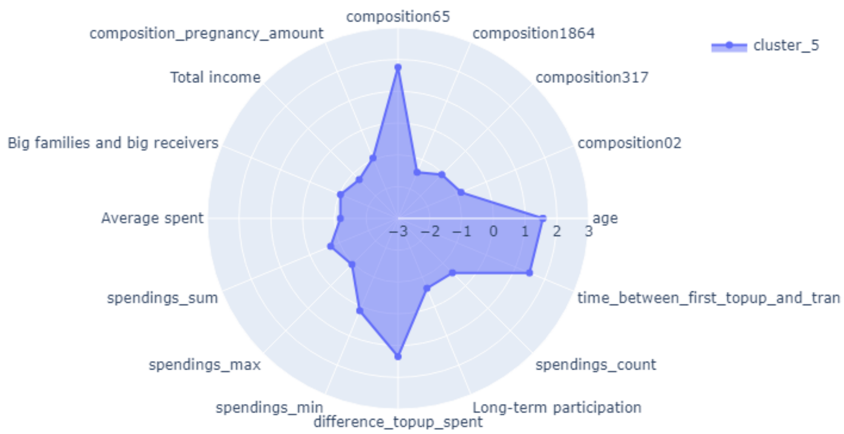


Figure 7.10: Radarplot of cluster 5

On the question if people from cluster 5 received AOV (government pension), a high fraction (40%) answered positive. Considering the relative high age of main beneficiaries in cluster 5, this is convincing. Another straightforward result from the cluster analysis on variables not included as input variable in K-Prototype, was about the type of communication channel. People were asked how they wanted to communicate with the staff of the Red Cross. People from cluster 5 had the smallest probability (6%) of using Whatsapp compared to the population (10%).

Another variable not used before was on how many people in the household were able to go to the supermarket. This was relatively low in cluster 4 and 5 compared to the other clusters. One of the reasons for this could be the relatively high age of people for cluster 5. But this is probably not the only reason, since cluster 4 consists of people with a lower age. Another complementary reason could be the low number of people in the household in cluster 5 compared to cluster 4. People from cluster 4 and 5 could have experienced more difficulties going to use their vouchers since they have less people in the household who are able to go to the supermarket.

7.4. Comparison with survey data

Although the survey data cannot be directly linked to the other data, we can still try to find similar groups in both datasets. The overlapping variables in the registration and survey data are the variables on age of the main beneficiary and the different household compositions. Using a frequency distribution of the clusters we can try to select only the relevant characteristics to be able to compare both datasets. Because these datasets are not possible to perfectly match, this section should be viewed as exploratory and not indisputable. The reason for still including the attempt for matching both datasets is twofold. One argument is to show that linking survey data with registration data is crucial during the setup of the project. The second argument is that this method can already give some indications of insights in clusters, but should not be considered definitive conclusions.

7.4.1. Method

In this case, the age of the main beneficiary can be visualized in figure 7.13. Cluster 5 with older people can be fairly easily separated since they are the only group with the bulk of their main beneficiaries having an age over 60. When also selecting the main beneficiaries with an age above 60 in the survey data, we could be able to get some insights on the needs and preferences of this cluster.

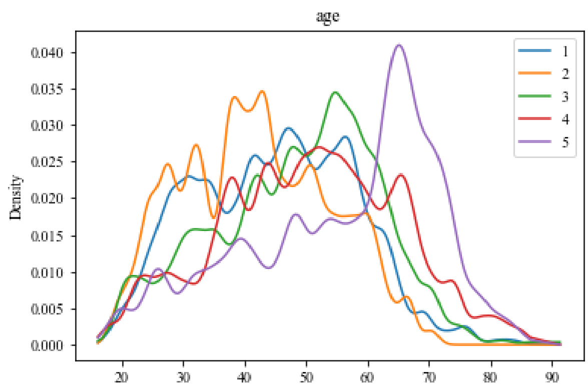


Figure 7.11: Density of age per cluster

To explain the procedure an example of a selection of subsets is shown in figure 7.12. In this example, all the recipients between 60 and 80 years old were considered. For each recipient in cluster 1, 2, 3 and 4 between these ages, it has been counted how many recipients are in this subgroup. These summations of all the people not in cluster 5 are visualized by the gray area in the graph. Also, the number of people in cluster 5 between 60 and 80 years old is calculated and can be seen by the sum of the grey and purple areas in the graph. The ratio between people in cluster 5 and between all the people between 60 and 80 years old has been calculated. In this specific example, the ratio of cluster 5 in the selection is 46%. The goal is to find a subset which maximizes this ratio, to be able to say that we have found a subset consisting mainly of people from cluster 5. This can be done by simulating many different subsets and comparing the found ratios of people from cluster 5.

This process was not only applied on the variable age, but also on all the other variables overlapping between both datasets. To maximize the potential of this process, also combinations between variables were used to find higher ratios. For example, we look for people between 60 and 80 years old, while at the same time looking for households consisting of more than 2 people above 65 years old. This section only articulates cluster 5 and no other clusters. The simulations of subsets also took other clusters into account as starting point, but cluster 5 seems to be the most promising to use, because of a relatively high number of outliers on age. This is useful since these are already somewhat divided by people in cluster 5 versus people not in cluster 5.

With all these simulations, we found a subset with the highest possible ratio. The conditions for this highest ratio is to have an age higher or equal to 69 and with 0 family members between 18 and 64 years old. The percentage of cluster 5 in this subset is 82.47. The remaining 17.53 percent is roughly evenly distributed among the four other clusters.

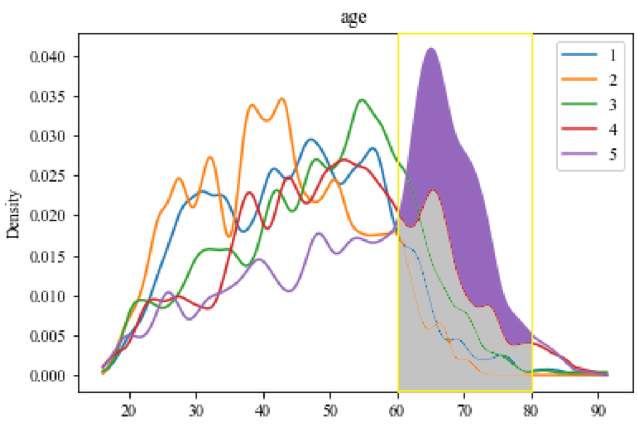


Figure 7.12: Density of age per cluster with an example selection

Since we know that the survey data is a subset of the registration/clustered data, we can try to find the subset with the highest ratio in the survey data. When applying the same conditions on the survey data, we find a subset in the survey data. Unfortunately, this is not a statistically sound method since there is still a 17.53 percent chance of selecting the wrong recipient. Also, there have not been statistical tests conducted on each variable, so we do not know if the subset consists of non-representative values on other variables. At this moment, no biases have been taken into account when comparing data. This should be done more extensively in future projects. But if we still want to say something about the subset in the survey data, we should do it with much care and it should be known that there are many underlying assumptions.

7.4.2. Results

Responses in the survey from cluster 5 were generally more positive compared to the rest of the population. Only 4.5 percent of the people in cluster 5 said that the evouchers were insufficient in fulfilling their needs, compared to 8.1 percent of the other groups. Furthermore, no one in this cluster had difficulties using the evouchers in the supermarket, compared to 3.7 percent of the other groups. Cluster 5 rated the communication and quality of distribution less positively compared to other groups. No one in cluster 5 experiencing corruption compared to others were 1.3 percent experienced corruption.

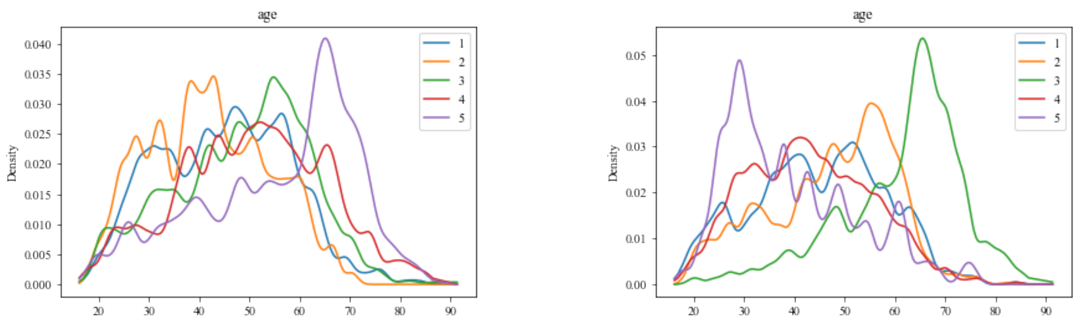
The findings of this survey comparison can be synthesized with the insights from the cluster analysis where we discovered that cluster 5 had a relatively long waiting time before they spend the first voucher in the supermarket. This is unlikely to have anything to do with having difficulties using the voucher, since we learned from the survey that cluster 5 is less likely to have problems using the vouchers. This could potentially be caused by the communication and/or quality of distribution where cluster 5 had more negative perceptions than other clusters. The cluster analysis also showed a relatively big difference in vouchers received and vouchers spent. Households from cluster 5 did not spend all their vouchers, which is remarkable given their relatively low income and given their response of having relatively few difficulties using these vouchers.

As stated before, these results need more validation and verification to be able to make strong statements. At the moment, these results are mainly used for demonstrating the ability to derive useful insights when synthesizing clustered and survey data. These results point to usefulness of using complaints and satisfaction of recipients, combined with their behavior.

7.5. Comparison with k-means

Since the cluster analysis showed that there were not many categorical variables of influence, a cluster analysis using the k-means algorithm has also been conducted. K-prototype consists of a combination of k-means and k-modes, so we only use the k-means part in this section. The same number of clusters (5) has been chosen to be able to compare the outcomes of k-prototype with k-means. Future research could analyse k-means on more number of clusters to find new undetected clusters. The silhouette coefficient for k-means, using 5 clusters, is 0.13 which is quite low. The consequence of this is that the cluster result of k-means can be called less compact compared to k-prototype. Therefore, the Davies-Bouldin (DB) index has also been consulted. This gives an indication of the the relative similarity of each cluster to each most similar cluster. Clusters with larger distances between each other, and smaller distances within their own cluster have a higher DB score, relative to other clusters. Using k-means with 5 clusters, the DB score is 1.86 which is not significantly different compared with using 2 to 9 clusters in k-means where DB scores varies between 1.71 and 2.19. Therefore, we can not conclude that this cluster result can be called robust.

When comparing both cluster results from k-means and k-prototypes, the following similarities and differences are remarkable and will be explained: age, sum of spending and the factor long-term participation are distinctly different. Figure 7.13 shows the density plot of age for k-prototype and for k-means. It can be concluded that k-means is more distinctively clustered on the variable age, since there are more peaks noticeable. We should be aware that the colors of both plots do not match since these clusters itself can not be perfectly compared. For example in this case, we see that cluster 5 using k-prototype (purple) and cluster 3 using k-means (green) show many similarities in being an outlier in the variable age. These recipients could be similar in both clusters, but this can not be said with high confidence and certainly not for the other clusters. Therefore, this comparison takes only the shape of different cluster variables into account and not their actual similarity. More research could benefit from a more in-depth comparison with different measures and tests of similarity.

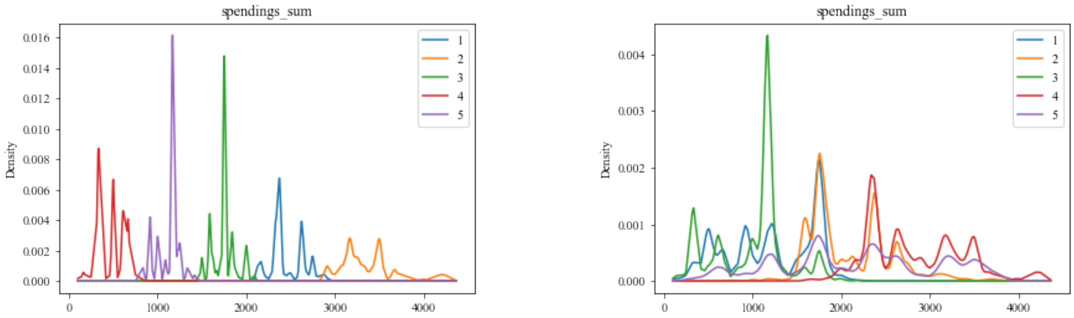


(a) Using k-prototype

(b) Using k-means

Figure 7.13: Age density plot of clusters

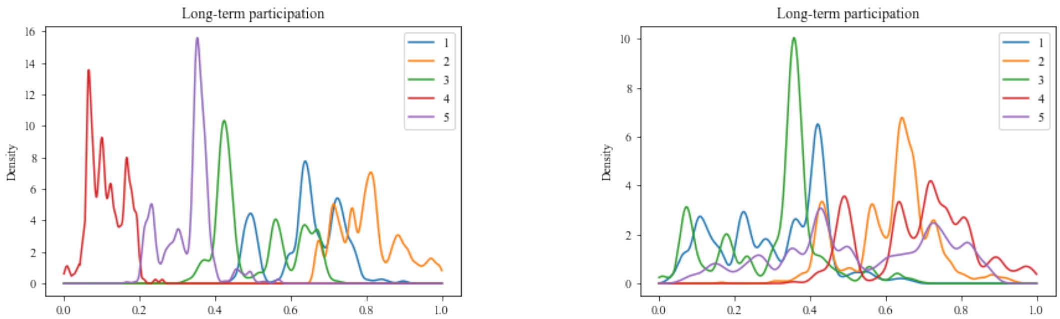
While age is more distinctively clustered using k-means, the variables on sum of spendings and on the factor 'Long-term participation' are less distinctively clustered using k-means. Figure 7.14 and figure 7.15 show that k-prototype clusters very well on these variables, while k-means struggles to find distinctive clusters using this variable. This is peculiar since the k-means clustering in this case uses less variables as input and can therefore have stronger cluster results, because there are less variables to take into account to find an optimum. But there could be a cleared distinction between clusters if more or less than 5 clusters are considered. This falls outside the scope of this research, but could be further analysed using multiple clusters.



(a) Using k-prototype

(b) Using k-means

Figure 7.14: Sum of spendings density plot of clusters



(a) Using k-prototype

(b) Using k-means

Figure 7.15: Long-term participation density plot of clusters

The conclusion on the comparison with cluster results with k-means is that the k-prototypes algorithm sometimes separates recipients more distinctively and in other times less. This is interesting considering the assumption that k-means with fewer variables would have a higher chance of creating stronger clusters. Although, the limitation of this is that only 5 clusters were considered for k-means. A different number of clusters would create different outcomes and could still theoretically outperform the k-prototype algorithm when looking at validation indices.

8

Discussion and Conclusion

This chapter discusses the results from chapter 7 and states the limitations and assumptions of these results. Furthermore, the academic and societal implications of these results are explained to get a clear overview of the added value of these results. Subsequently, the main research question will be answered and the link with the EPA (Engineering & Policy Analysis) curriculum will be given. This research finishes with recommendations for humanitarian aid workers and for academics. These recommendations will be focused on improving current practices and on possibilities for future research.

The following subquestions are being answered in this chapter:

- What are the limitations of this study and how are they affecting the outcome?
- How can the outcomes of this study be generalised to crises where CVA projects are used?
- How can recipients of cash and voucher assistance be categorized using the field of consumer segmentation by using machine learning methods?
- What are recommendations for humanitarian aid workers and academics?

8.1. Discussion

8.1.1. Interpretations of results

In this research, theory from customer segmentation literature was used to come up with a clustering algorithm and variables on customer behaviour (the RFM model). The following findings were obtained from the results of the Sint Maarten case study:

- The five different distinct clusters were found using the geo-demographic data and behaviour data.
- The geo-demographic variables that were best in characterizing clusters are: the age of the main beneficiary, the different household compositions and the factor 'big families and big receivers'.
- The most distinguishable behaviour variables were: the number of supermarket visits (frequency), the time between when the first voucher received and the first transaction (recency) and the variables on the amount of money that was spent with the vouchers (monetary).
- It was found that the variables gender and 'the most frequent time of day of visiting a supermarket' are not effective in explaining characteristics of clusters, using this specific case.
- To include the 'voice' of recipients (needs and preferences) in the model, a connection between the registration data & behavioural data and the survey data is needed.A common identifier to make the connection was not available in this case.
- Comparing a cluster (5) with the survey data gives an indication of receiving more insights in the needs, preferences and behaviour of recipients.
- When using the same number of clusters and input variables in a numerical-only clustering algorithm (k-means), the cluster result does not necessarily lead to stronger separated clusters.
- Insight in recipients can be used for retargeting recipients of CVA using a feedback loop when CVA projects are prolonged.

8.1.2. Limitations

This research study has several limitations. This section elaborates on the main 7 limitations.

The cluster assumption

The first limitation lies in the method itself: clustering and factor analysis assume that there is an underlying structure existing of groups. An inherent assumption with clustering is that it is impossible to statistically validate clusters due to the aspect of unsupervised learning. There are no labels of the clustered groups beforehand to validate the findings, so real validation with test and train data is not possible. A possible next step in research could be to include the use of classification models where data with labels is used to validate clusters. The limitation of these assumptions lie in the fact that clustering is to some extent a subjective field of research. Therefore, the findings are less likely to be directly generalizable to other cases since findings differ due to the subjective nature of choosing relevant variables and the number of clusters. But the method itself could be regarded useful for practitioners in the field since they gain more insight in the behaviour and geo-demographic variables.

Limitations in the use of survey data

The survey data is an important aspect by which the 'voice' of recipients, such as the presence of needs, can be incorporated in CVA projects. However the use of survey data in this case was limited due to the lack of a common identifier between the survey data and the other data sets on registrations and behaviour. The registration data and behaviour data sets were interlinkable and therefore it was possible to combine the variables of these data sets in one clustering model. Because the survey data was not interlinkable with a common identifier, this data was not used in the clustering data itself, it was only used for the validation and analysis of the clusters. Additionally, for the validation and analysis of the clusters, the survey data could only be used for clusters showing a distinguishable variable that was overlapping in survey and clustered data, but this has many implications and is not a credible method.

Validation Indices

More internal cluster validation indices could have been used to compare the internal structures of clusters. This research only uses the Silhouette Index to get insight in how well each point is matched to its own cluster and how poorly it matches with neighboring clusters. To improve the evaluation of the clusters it is recommended to apply additional validation methods. For example, The Dunn Index, Gamma Index and G+ Index can also be utilized for an increased understanding in the cohesion, separation and compactness of the cluster result. This research was limited by time and computing power to run the calculations of these heavy algorithms. Future research could focus on the variables that are generating more cohesive, separated and compact clusters, based on various indices. Other validation measures could be performed as well. For example, the stability of the found clusters can be explored to get insight in the influence of removing recipients on the cluster outcomes. This simulation is also very computationally expensive and requests certain knowledge on nondeterministic polynomial time (NP) in computational complexity theory. The cluster results could benefit from these validation measures since this could make the clusters more robust which provides additional proof on the applicability into the real world. This limitation has been mitigated to some extent in this research by inspecting and removing outliers before clustering, to make the cluster model less sensitive for extreme values.

Disadvantages of the use of K-prototype

Another limitation comes from the selection of the cluster algorithm. This research uses an algorithm called K-prototype, which is useful for its applicability on mixed data (both nominal and ordinal data) but also has disadvantages. One disadvantage is the speed and possibly accuracy compared to only-numerical algorithms. But many datasets consist of nominal and ordinal data, and especially when considering data on needs in the humanitarian sector, the presence of mixed data can not be ignored. An alternative could be to use a numerical algorithm (e.g., k-means), but with one-hot encoded categorical variables. This however increases the dimensionality of the data even more and is therefore not preferred. Another limitation on K-prototype is the parameter selection since a weighting called γ balances categorical and numerical data. Different weightings of variables can steer cluster outcomes to your direction of interest. In this research a standard γ has been chosen which balances numerical and categorical variables, creating as much as a neutral starting point of the cluster algorithm. This research compared the results of 5 clusters with a similar exercise using k-means. The conclusion from this comparison is that k-means does not specifically create stronger clusters (considering the same number of clusters as used in k-prototype) because the average Silhouette coefficient was lower than the one used in k-prototype. More research using different numbers of clusters could lead to more convincing insights in why and when k-means could outperform k-prototype.

The curse of high dimensionality

Working with the curse of high dimensionality gives challenges, since it is unclear when results are negatively influenced by this phenomenon. This is dependent on the specific domain, the number of observations and variables, but also on the correlation between those variables. There are no rules of thumb or mathematical proof for getting an indication of the appropriate number of input variables for the clustering algorithm. Therefore, multiple clustering subsets were being run and numerous outcomes were manually analysed. This manual inspection of clusters is subjective to a certain extent and therefore has limitations to identically reproduce.

Results specific to the Sint Maarten case

Using the specific case of Sint Maarten, the context and data cause a number of limitations. Although Sint Maarten was devastated by hurricane Irma and Covid-19, the citizens are considered relatively wealthy compared to other countries where the Red Cross has active CVA programmes. The relatively high income can influence perceptions of the evoucher project since citizens are already accustomed to a high consumption level. Satisfaction levels and perceived needs could therefore be less articulating since this was not considered a project with many different problems or needs. Additionally, Sint Maarten is a small and very dense island, comparable to a large city. Therefore, the geographic variables on districts and neighbourhood were not very distinctive since everything is quite close to each other. CVA projects in rural areas are expected to have more explanatory value on geographic variables. More research on different cases in the various regions of the world would benefit for making this more applicable to a general new case.

Non-applicability to cash in hand and short projects

Another inherent limitation by using evouchers is that this will not be applicable on projects where cash in hand is the chosen delivery mechanism. This has to do with data availability since cash in hand is not possible to track due to the absence of registered transactions, which is used for behavioural data. It is estimated that a large portion of CVA makes use of this delivery mechanism, making it inappropriate for application on a large portion of CVA projects. Nevertheless, this research is applicable for evouchers or trackable credit/debit cards. Since this research proposes the use of a feedback loop during a running CVA project, it is less applicable when this project has a short duration. Shorter duration means less chance of retargeting recipients of CVA and means less generated behaviour data. Therefore, short projects are less useful when applying the methods in this research.

8.1.3. Implications

What do the results implicate for the academic world and society? The contribution to these two fields will be explained here.

Academic Contribution

The contribution to the scientific literature by doing this research comes from two aspects. First of all, most of the existing literature consists of papers written by humanitarian organisations and of consultancy firms doing evaluations on past projects. Most of this literature is grey and the scientific base is sometimes lacking (Obrecht, 2017). A lot of qualitative research has been done, but the quantitative research is lagging behind due to a lack of data (High Level Panel on Humanitarian Cash Transfers, 2015). This research uses a data-driven approach with a novel method of clustering recipients. This exploration proposes a new methodology to capture dynamic behaviour of recipients of evouchers using clustering on voucher data. This is the first study that combines RFM and clustering to reveal trends in behaviour using humanitarian aid.

Getting back to the knowledge gaps from the literature: this research proves to be able to include the impact on and the satisfaction of recipients of evouchers. Also, different household compositions and recipient characteristics can be selected and analyzed using the proposed factor-cluster approach. The third knowledge gap on including ways to capture recipient behaviour is only partly filled since this study did not particularly focus on identifying behaviour such as saving, panic-buying and selling. However, it has been identified that these behaviours can be further investigated using time-series analysis combined with clustering.

Societal Contribution
One could argue that humanitarian aid (and research on improving this) inevitably contributes to society. Helping people in need does not need more explanation to prove a societal contribution. This research however, contributes to bringing a new method to humanitarian aid to evaluate impact and to assist in the decision-making process of humanitarian aid workers. Therefore it is only indirectly contributing to humanitarian aid and society. This method can be used as new input for targeting methods to better incorporate the needs of users of evoucher and ecash projects. The feedback loop proposed in figure 2.3 can be further studied to assist recipients with greater efficiency and effectiveness with the use of behaviour data and survey data on needs. The cluster outcomes from this study can be applied by humanitarian aid workers on new variables in, for example, proxy means tests and other methods of targeting. Also, humanitarians working on human-centered design can use these cluster results to confirm or enrich their often used personas of recipients of CVA. For humanitarian aid workers, this research mostly contributes to a new method of receiving insights in recipients.

8.2. Conclusion

This section consists of the conclusion answering the main research question and on the link of this research to the curriculum of Engineering, Policy and Analysis.

8.2.1. Answering the main research question

Main research question:

How can recipients of cash and voucher assistance be categorized using the field of consumer segmentation by using machine learning methods?

Recipients of cash and voucher assistance can be categorized using a two-step data modelling approach. This consists of dimension reduction on variables with factor analysis followed by a clustering algorithm to find distinguishable segments. It is beneficial to use three sources of data as input for categorizing recipients: geo-demographic data, behavioural data using the RFM model and survey data on the satisfaction and needs of voucher recipients. These data sources are required to be linked to get the most insights out of the cluster analysis. For future CVA projects, the found subgroups can be used for targeting by building more accurate personas and by further research on identifying new recipients into existing clusters.

It is unknown if all clusters results are always relevant since not all clusters show obvious patterns. Nonetheless, this research shows that cluster 5, when provisionally combined with data on preferences and problems experienced, could give meaningful insights in the voices of recipients. Although, this conclusion should be taken with caution since survey data can not be perfectly matched with behaviour data,due to the lack of a common identifier..

This novel research proposes a machine learning workflow which aims to quickly cluster recipients of evouchers. With more research on the limitations, this workflow can be generalized to be used by others on similar CVA projects. This analysis shows how (often unused) data sources and machine learning models can be used to improve the categorization of recipients and give new insights to the humanitarian field. It could potentially be deployed quickly and responsively on CVA projects where cash is being tracked (ecash and evouchers).

8.2.2. Link to the EPA program

Since this Master Thesis is part of the Engineering and Policy Analysis (EPA) master program of the TU Delft, this Master Thesis supports the criteria given by EPA2942, the Master Thesis course of EPA. This entails that the work should contain an analytical component, is multidisciplinary in nature, focuses on a technical application or domain and relates to a grand challenge according to Studiegids. This research includes a so-called grand challenge,namely improving targeting in cash and voucher assistance aid projects, and aims to provide local decision-makers and humanitarian workers with new insights in the results and working mechanisms of Cash and Voucher Assistance. This research found that with data-driven insights using clustering recipients, decision-makers can be supported in solving complex socio-technical problems. Therefore, it can be concluded that this research is conducted according to the requirements of the EPA master program.

8.3. Recommendations

Several recommendations can be directly or indirectly derived from this research. Recommendation can be categorized in future research for academics and in implementations for humanitarian aid workers.

8.3.1. For academics

Recommendations are constructed with the following questions in mind: what if this research continued for three more months and what would be the first thing to recommend?

Validation

An element of this research which should be done more extensively is the validation of cluster performance. This study only uses expert validation and the Silhouette Index to evaluate the cluster result. It would be beneficial to dive more into other indices such as the Davies-Bouldin, Gamma, Dunn and G+ Indices. These indices will give more evidence on the internal compactness and cohesion of clusters and will therefore contribute to the degree of applicability of this method. It could be, for example, that one of the cluster results scores poorly on a specific index, indicating that these findings are not very robust and should be interpreted with care. Implementation and more research on cluster validation will solve this challenge.

High dimensionality

Another missing piece of the puzzle is to be able to check if a cluster model is built for high dimensionality. At this moment, it was actively tried to reduce the number of variables to make the clustering result less exposed to the curse of high dimensionality. However, the curse of high dimensionality is not convincingly being diminished. Metrics on suffering from the curse of high dimensionality should be researched to get more insight in the influence of high dimensionality on cluster results. This can be done by running more simulations of different subsets of variables and by comparing cluster results and their validation indices. Another recommendation for further research in line with the curse of high dimensionality is to conduct research on sufficient sample size for clustering.

Generalization to other cases

This novel method for categorizing recipients of CVA requires careful examination of variables and is therefore not directly generalizable to other cases, especially not when being used by humanitarian aid workers without understanding of the limitations. Therefore, this modelling pipeline should be made more stream-lined and further substantiated. Academics can conduct research on ways to make a useful tool out of this method, considering the lack of background knowledge of workers in the field. Communication on the limitations is therefore much needed when further developing a tool for clustering. Another important element which should be analyzed more is following the same procedures of this approach, but with more data using a variety of cases. This could demonstrate the contextual dependence of these models.

Choices between mixed, only numerical or categorical data

Although categorical clustering has been argued as suitable for this research, this could be further examined by comparison of different algorithms used for numerical or only categorical variables. The motivation for k-prototype can be additionally validated by this. If this is not possible, minimization or maximization of the categorical variables and minimization or maximization of γ using the k-prototype should be tested in multiple simulations. Other considered cluster algorithms worth mentioning for further research: hierarchical clustering (agglomerative and complete linkage), density based clustering (MULIC, CLIQUE, HIERDENC) or model based clustering.

Future possibilities enabled by this research

The result of this research can also be regarded as typologies which are contextual ready and a new insightful method of revealing patterns and understanding needs. This behaviour data can be used in

different other studies on CVA, for example when studying the use of agent-based models (ABM) in CVA. When further researched, these clusters can be used as a narrative for the construction of ABM. The effect of policies can be simulated using these kinds of models, providing insights for decision-makers on the effect of policies on certain clusters.
Other follow-up studies can dive into market basket analysis and time series analysis of transactions using the RFM model, as already proposed in chapter 5. Another potentially beneficial research which can be conducted with the found results is prediction. For example, it can be assumed that the found clusters are static and that new recipients are part of one of the clusters. A prediction model can be made to predict to which cluster the new recipient belongs.

8.3.2. For humanitarian aid workers

Recommendations for humanitarian aid workers are composed of the perception of actually working and implementing improvements in the field of CVA. It is important to try to think more steps ahead when recommending policies for humanitarian aid workers, since CVA is part of a bigger system of multiple actors and various dependencies.

Implementation

Since this research brings new insights which can be used in targeting methods, further implementation could be considered on how to consolidate these insights. A few recommendations of absorbing these insights into daily practises are to make more use of feedback loops during CVA projects. The findings of this research could be implemented as follows: usually a CVA project only monitors satisfaction and needs at the beginning and at the end of a project, but this research argues that it would be beneficial to conduct intermediate monitoring surveys and recipient analyses and use them for interim evaluation of the project. This is currently not done structurally in a CVA project, but only occasionally when projects are prolonged. Since research points to including recipient voices more, a straightforward recommendation is to actually collect this feedback more often. This can be done with, for example, more frequent surveys on a representative sample of the total population receiving CVA. With this, a monthly factor-cluster analysis using behavioural data and survey data can be conducted. A quick analysis by humanitarian aid workers on the generated clusters can lead to an informed overview of how the project is running. If necessary, aid workers can implement additional policies, as outlined in chapter 2, to assist clusters with special needs.
For example, it could be discovered that a certain cluster is experiencing problems with the communication on CVA and that this same group pointed out that the majority does not speak English (or any other common language) while living at a certain district. This group can be assisted when focusing more on delivering communication in a convenient language in the specific district. Another hypothetical situation could be when a cluster did not use their vouchers as often as other clusters, while at the same time responding in the survey that they are satisfied with the amount of vouchers given. This could indicate to this cluster being the least in need compared to other clusters. After further confirmation using more in-depth analysis and possibly interviews, this could lead to lowering the transfer value for this specific cluster and raising the transfer value for the people more in need. Although this could sound controversial, in the humanitarian sector there are limited resources and the goal is to alleviate suffering and bring the population back to pre-crisis level as soon as possible. This can be achieved by more efficient resource allocation using the proposed feedback loop in figure 8.1.

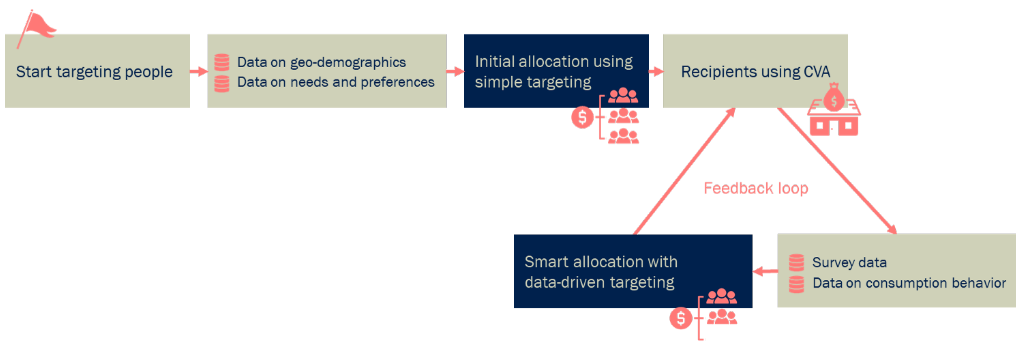


Figure 8.1: New retargeting flow with feedback loop

Data minimization
A practical disadvantage of this targeting method is that conducting surveys take time, both for recipients as for aid workers. Therefore, more thoughts should be given to the design of surveys. One of the side effects of this research is knowledge on which data is not distinguishable in the cluster result, indicating a low degree of explanatory value. Therefore, data minimization can be applied in the design of a survey. This research signals that the variable gender was not of use in describing any of the clusters. Although, this can not be directly extrapolated on other cases, this indicates that not all the data is equally useful in this factor-cluster approach. More research on data minimization can lead to stronger arguments of including or excluding certain often asked questions in surveys. A byproduct of this research is knowledge if this method can be used in data-scarce settings, because many cases in the humanitarian sector are data scarce. A straightforward recommendation on survey design is to link survey data with registration data with a common identifier. This makes segmentation on needs much easier for future projects.

Survey design
One strong recommendation for designing surveys is to add triangulation questions when asking data on satisfaction and needs. Although this will increase the length of the survey, it can be beneficial to get a clear as possible insight of the situation of the recipient. For example, if surveys include only one question on to which extent the voucher was sufficient, recipients might be inclined to give a favourable answer which does not fully reflect their real satisfaction. For that reason, triangulation questions could be used to include multiple questions where the outcomes are indicative for one variable measured. For instance, an additional question could be to ask if they could have bought everything they needed in the time after they received their vouchers. Another additional question for triangulation on this could be to ask if they worry often about if they can buy everything they need by using the voucher. The combination of these questions could lead to a more validated survey response on to which extent the voucher was sufficient. A related recommendation is to include five point scale on these survey answers, which is an improvement of the three point scale used in the case of Sint Maarten. A larger range of answer values makes it possible to transform the variables from ordinal to interval scale later in the data modelling process. Another practical recommendation for survey design which was discovered by this research is to try to prevent open questions. The only practical way to include open questions is to place them at the end of a specific funnel of closed questions. In this specific case, data analysis became increasingly hard when people used open questions to make a large variety of needs clear in six different languages.

Development of a tool for in the field
As already mentioned as recommendation for academics, a tool can be made on the application of factor-cluster approach. Although, academics should dive more into the treatment of limitations of this model, humanitarian aid workers can already start with thinking about what they want to include in such a tool. For example, visualizations of clusters on geographic areas can be immensely useful when responsibilities of an evoucher project are divided over different areas. The design of a dashboard of this factor-cluster tool should be made by people working in the field to maximize the usefulness. Human-centered design departments of humanitarian organisations are looking for ways to connect quantitative part of projects with qualitative projects but do not know how. This user-centred quantitative research could help.

Minimizing targeting errors
Humanitarian aid workers can conduct research on targeting errors. It is proposed that with this research, targeting errors can be further minimized. This should be assessed by using control groups in various dissimilar cases for delivering the best evidence of improving targeting by reducing errors. One limitation of this recommendation is that the usefulness of the new targeting method strongly depends on the length of the project. For a short project this might not work, since not enough feedback loops can be passed and not enough useful behaviour data can be collected. These recommendations will hopefully contribute to a more user-centered approach of cash and voucher assistance, so that people in need can be better heard and more effectively assisted.

Bibliography

510. (2016). 510 Mission & Vision. https://www.510.global/510-mission-vision/. (Cit. on p. 30)

510. (2018). *Data responsibility policy* (tech. rep.). Netherlands Red Cross. (Cit. on p. 22).

Aaker, J. L., & Fournier, S. (1995). A brand as a character, a partner and a person: Three perspectives on the question of brand personality. *ACR North American Advances* (cit. on p. 41).

Abbasimehr, H., & Shabani, M. (2019). A new methodology for customer behavior analysis using time series clustering: A case study on a bank's customers. *Kybernetes* (cit. on p. 43).

Aggelis, V., & Christodoulakis, D. (2005). Customer clustering using rfm analysis. *Proceedings of the 9th WSEAS International Conference on Computers*, 2 (cit. on p. 44).

Agrawal, R., Imielinski, T., & Swami, A. N. (1993). Mining association rules between sets of items in large databases. *SIGMOD '93* (cit. on p. 43).

Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using market basket analysis in management research. *Journal of Management*, *39*(7), 1799–1824 (cit. on p. 43).

Ahmad, A., & Khan, S. S. (2019). Survey of state-of-the-art mixed data clustering algorithms. *IEEE Access*, *7*, 31883–31902 (cit. on p. 52).

Aiken, E., Bellue, S., Karlan, D., Udry, C. R., & Blumenstock, J. (2021). *Machine learning and mobile phone data can improve the targeting of humanitarian assistance* (tech. rep.). National Bureau of Economic Research. (Cit. on pp. v, 9).

Akay, Ö., & Yüksel, G. (2018). Clustering the mixed panel dataset using gower's distance and k-prototypes algorithms. *Communications in Statistics-Simulation and Computation*, *47*(10), 3031–3041 (cit. on p. 52).

Altındağ, O., O'Connell, S. D., Şaşmaz, A., Balcıoğlu, Z., Cadoni, P., Jerneck, M., & Foong, A. K. (2021). Targeting humanitarian aid using administrative data: Model design and validation. *Journal of Development Economics*, *148*, 102564 (cit. on pp. 3, 20).

Ark-Yıldırım, C., & Smyrl, M. (2021). *Social Cash Transfer in Turkey: Toward Market Citizenship* (1st ed. 2021). Palgrave Macmillan. (Cit. on p. 22).

Aschenbruck, R., & Szepannek, G. (2020). Cluster validation for mixed-type data. *Archives of Data Science, Series A*, *6*(1), 02 (cit. on p. 54).

Bailey, S. (2016). *Why not cash? The case for cash transfers for refugees in Mozambique* (tech. rep.). Overseas Development Institute. (Cit. on p. 2).

Barrett, C. B., Lentz, E., & Burton, L. (2013). Hunger and food insecurity. (Cit. on p. 3).

Basurto, M. P., Dupas, P., & Robinson, J. (2020). Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural malawi. *Journal of public economics*, *185*, 104047 (cit. on p. 20).

Bellman, R. (1961). Curse of dimensionality. *Adaptive control processes: a guided tour. Princeton, NJ*, *3*(2) (cit. on p. 47).

Benhassine, N., Devoto, F., Duflo, E., Dupas, P., & Pouliquen, V. (2015). Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education. *American Economic Journal: Economic Policy*, *7*(3), 86–125. https://doi.org/10.1257/pol.20130225 (cit. on p. 21)

Bijmolt, T. H., Paas, L. J., & Vermunt, J. K. (2004). Country and consumer segmentation: Multi-level latent class analysis of financial product ownership. *International Journal of Research in Marketing*, *21*(4), 323–340 (cit. on p. 42).

Birant, D. (2011). Data mining using rfm analysis. *Knowledge-oriented applications in data mining.* IntechOpen. (Cit. on p. 44).

Browne, C., Matteson, D. S., McBride, L., Hu, L., Liu, Y., Sun, Y., Wen, J., & Barrett, C. B. (2021). Multivariate random forest prediction of poverty and malnutrition prevalence. *PloS one*, *16*(9), e0255519 (cit. on p. 3).

CaLP. (2018). *State of the World's Cash Report 2018* (tech. rep.). The Cash Learning Partnership. (Cit. on pp. 4, 21).

CaLP, & IARAN. (2019). *The Future of Financial Assistance: An outlook to 2030.* (tech. rep.). (Cit. on pp. 2, 3).

CARE International, & ideas42. (2020). *Applying Behavioral Science to Humanitarian Cash & Voucher Assistance for Better Outcomes for Women.* (tech. rep.). (Cit. on pp. 4, 23, 55, 56).

Cash Learning Parthership. (2020). *Winds of Change: Lessons and Recommendations on the Use of Cash and Voucher Assistance (CVA) for the Caribbean Atlantic Hurricane Season.* (tech. rep.). CaLP. (Cit. on pp. 8, 28).

Cash Learning Partnership. (2020a). *Glossary of terminology for cash and voucher assistance* (tech. rep.). CaLP. (Cit. on pp. 15, 17).

Cash Learning Partnership. (2020b). *State of the World's Cash Report 2020* (tech. rep.). Cash Learning Partnership. (Cit. on pp. v, 2, 3, 5).

Castillo, J. G. (2021). Deciding between cash-based and in-kind distributions during humanitarian emergencies. *Journal of Humanitarian Logistics and Supply Chain Management* (cit. on p. 4).

CBS News. (2017). Third of buildings on Dutch St. Martin destroyed. (Cit. on p. 26).

Chen, D., Guo, K., & Ubakanma, G. (2015). Predicting customer profitability over time based on rfm time series. *International Journal of Business Forecasting and Marketing Intelligence*, *2*, 1. https://doi.org/10.1504/IJBFMI.2015.075325 (cit. on p. 44)

Chen, Y., Rajabifard, A., Sabri, S., Potts, K. E., Laylavi, F., Xie, Y., & Zhang, Y. (2020). A discussion of irrational stockpiling behaviour during crisis. *Journal of Safety Science and Resilience*, *1*(1), 57–58. https://doi.org/10.1016/j.jnlssr.2020.06.003 (cit. on p. 5)

Chouikhi, H., Charrad, M., & Ghazzali, N. (2015). A comparison study of clustering validity indices. *2015 global summit on Computer & information technology (GSCIT)*, 1–4 (cit. on p. 54).

CIA. (2021). Sint Maarten. (Cit. on pp. 25, 26).

Cleveland, M., Babin, B. J., Laroche, M., Ward, P., & Bergeron, J. (2003). Information search patterns for gift purchases: A cross-national examination of gender differences. *Journal of Consumer Behaviour: An International Research Review*, *3*(1), 20–47 (cit. on p. 42).

Connected Papers. (2022). Connected Papers - Explore connected papers in a visual graph. (Cit. on p. 15).

Creusen, M. E. (2010). The importance of product aspects in choice: The influence of demographic characteristics. *Journal of Consumer Marketing* (cit. on p. 42).

Croft, M. J. (1994). *Market segmentation: A step-by-step guide to profitable new business.* Cengage Learning Emea. (Cit. on p. 41).

Cuevas, P., Kaan, I., Twose, A., & Çelik, Ç. (2019). *Vulnerability and Protection of Refugees in Turkey: Findings from the Rollout of the Largest Humanitarian Cash Assistance Program in the World* (tech. rep.). World Bank and World Food Programme. Washington, DC. (Cit. on p. 2).

Daellenbach, K., Parkinson, J., & Krisjanous, J. (2018). Just how prepared are you? an application of marketing segmentation and theory of planned behavior for disaster preparation. *Journal of nonprofit & public sector marketing*, *30*(4), 413–443 (cit. on p. 3).

Darcy, J., Stobaugh, H., Walker, P., & Maxwell, D. (2013). *The use of evidence in humanitarian decision making ACAPS operational learning paper.* (tech. rep.). Feinstein International Center and Tufts University. Boston. (Cit. on p. 3).

Darroch, G. (2017). Willem-Alexander: Sint-Maarten destruction 'worse than any war zone'. (Cit. on p. 26).

De Hamer, J. (2019). *Disaster governance on St. Maarten* (Doctoral dissertation). Wageningen University. Wageningen. (Cit. on p. 27).

Department of Statistics Sint Maarten. (2017). Statistical Yearbook 2017. (Cit. on pp. 33, 34).

Department of Statistics Sint Maarten. (2019). Labour Force Survey 2019. (Cit. on p. 34).

Development Initiatives. (2021a). *Development Initiatives.* (tech. rep.). Development Initiatives. (Cit. on p. 8).

Development Initiatives. (2021b). *Global Humanitarian Assistance Report 2021* (tech. rep.). Development Initiatives. (Cit. on pp. 2, 29).

Devidal, P. (2021). Cashless cash: financial inclusion or surveillance humanitarianism? https://blogs.icrc.org/law-and-policy/2021/03/02/cashless-cash/. (Cit. on p. 21)

Dolnicar, S., & Grun, B. (2011). Three good reasons not to use factor-cluster segmentation (cit. on p. 48).

Dolničar, S. (2004). Beyond "commonsense segmentation": A systematics of segmentation approaches in tourism. *Journal of Travel Research*, *42*(3), 244–250 (cit. on p. 41).

Drayer, D. (2016). Sint Maarten broeinest van mensenhandel en uitbuiting. (Cit. on p. 25).

Dutch Relief Alliance. (2018, October 1). *121: Improving cash-based assistance.* https://dutchrelief.org/121-improving-cash-based-assistance/. (Cit. on p. 21)

ECHO. (2013). *The use of cash and vouchers in humanitarian crises* (tech. rep.). European Commission. Brussel. (Cit. on p. 17).

ECHO. (2017). *Guidance to Partners Funded by the Directorate-General for European Civil Protection and Humanitarian Aid Operations (ECHO) to Deliver Large-Scale Cash Transfers.* (tech. rep.). European Commission. Brussel. (Cit. on p. 2).

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of management review*, *14*(4), 532–550 (cit. on p. 8).

Formann, A. K. (1984). *Die latent-class-analyse: Einführung in theorie und anwendung.* Beltz. (Cit. on p. 50).

Foss, A., Markatou, M., Ray, B., & Heching, A. (2016). A semiparametric method for clustering mixed data. *Machine Learning*, *105*(3), 419–458 (cit. on p. 54).

Frank, R., Massy, W., & Wind, Y. (1972). Market segmentation. (Cit. on p. 42).

Gelan, A. (2006). Cash or food aid? a general equilibrium analysis for ethiopia. *Development Policy Review*, *24*(5), 601–624 (cit. on p. 4).

Goldberg, D. E. (1989). Optimization, and machine learning. *Genetic algorithms in Search* (cit. on p. 53).

Government of Sint Maarten. (2017). *Sint Maarten National Recovery and Resilience Plan: A Roadmap to Building Back Better* (tech. rep.). Government of Sint Maarten. Philipsburg. (Cit. on p. 26).

Ground Truth Solutions. (2018). Participation revolution? (Cit. on p. 2).

Ground Truth Solutions. (2019). Changing the perspective: what recipients think of cash and voucher assistance. (Cit. on pp. 2, 21).

Hadi, N. U., Abdullah, N., & Sentosa, I. (2016). An easy approach to exploratory factor analysis: Marketing perspective. *Journal of Educational and Social Research*, *6*(1). https://doi.org/10.5901/jesr.2016.v6n1p215 (cit. on p. 50)

Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell system technical journal*, *29*(2), 147–160 (cit. on p. 53).

Hanna, R., & Olken, B. A. (2018). *Universal basic incomes vs. targeted transfers: Anti-poverty programs in developing countries* (Working Paper No. 24939). National Bureau of Economic Research. https://doi.org/10.3386/w24939. (Cit. on pp. 3, 20)

Harvey, P., & Pavanello, S. (2018). *Multi-purpose cash and sectorial outcomes, a review of evidence and learning.* UNHCR Global Cash Operations. (Cit. on p. 21).

Harvey, P., & Bailey, S. (2011). *Cash transfer programming in emergencies.* Humanitarian Practice Network, Overseas Development Institute. (Cit. on p. 17).

HelpAge International. (2010). *Cash transfers in emergencies: A practical field guide.* (tech. rep.). HelpAge International. Muang, Chiang Mai. (Cit. on p. 17).

Hennig, C., & Lin, C.-J. (2015). Flexible parametric bootstrap for testing homogeneity against clustering and assessing the number of clusters. *Statistics and Computing*, *25*(4), 821–833 (cit. on p. 53).

Hermans, B., & Kösters, L. (2019). *Labour force on the Dutch Caribbean islands* (tech. rep.). CBS. The Hague. (Cit. on p. 25).

High Level Panel on Humanitarian Cash Transfers. (2015). *Doing cash differently - How cash transfers can transform humanitarian aid* (tech. rep.). Overseas Development Institute. (Cit. on pp. 2, 4, 73).

Hu, Y.-H., & Yeh, T.-W. (2014). Discovering valuable frequent patterns based on rfm analysis without customer identification information. *Knowledge-Based Systems*, *61*, 76–88. https://doi.org/https://doi.org/10.1016/j.knosys.2014.02.009 (cit. on p. 43)

Huang, Z. (1997). Clustering large data sets with mixed numeric and categorical values. *Proceedings of the 1st pacific-asia conference on knowledge discovery and data mining,(PAKDD)*, 21–34 (cit. on pp. 52, 53).

Iam-On, N., Boongoen, T., & Kongkotchawan, N. (2014). A new link-based method to ensemble clustering and cancer microarray data analysis. *International Journal of Collaborative Intelligence*, *1*(1), 45–67 (cit. on p. 52).

ideas42. (2019). *Cash and Change Using Behavioral Insights to Improve Financial Health in Three Cash Transfer Programs* (tech. rep.). https://www.ideas42.org/wp-content/uploads/2019/09/I42-1160_CashTransfers_paper_final-4.pdf. (Cit. on p. 21)

IFC. (2018). *Kakuma as a Marketplace. A consumer and market study of a refugee camp and town in northwest Kenya.* (tech. rep.). International Finance Corporation, Washington DC. (Cit. on p. 5).

IFRC. (2020). *Global Plan 2021* (tech. rep.). International Federation of Red Cross and Red Crescent Societies. Geneve. (Cit. on p. 2).

IFRC. (2021). *Practical guidance for data protection in cash and voucher assistance* (tech. rep.). International Red Cross and Red Crescent Movement. (Cit. on p. 21).

International Monetary Fund. (2021). Curaçao and Sint Maarten: Staff Concluding Statement of the 2021 Article IV Mission. (Cit. on pp. 4, 8, 26).

International Red Cross and Red Crescent Movement. (2022). Cash in emergencies toolkit. https://cash-hub.org/guidance-and-tools/cash-in-emergencies-toolkit/. (Cit. on p. 17)

IPCC. (2021). *Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (tech. rep.). Cambridge University Press. Cambridge. (Cit. on pp. 1, 26).

IRC. (2016). *Cost Efficiency Analysis: Unconditional Cash Transfer Programs.* (tech. rep.). International Rescue Committee. (Cit. on p. 4).

Jadczaková, V. et al. (2013). Review of segmentation process in consumer markets. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *61*(4), 1215–1224 (cit. on pp. 42, 44).

Kaiser, H. F. (1974). An index of factorial simplicity. *psychometrika*, *39*(1), 31–36 (cit. on p. 50).

Key Aid Consulting. (2018). *Social transfers study in the Gaza Strip.* (tech. rep.). (Cit. on p. 5).

Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on rfm analysis of customer purchase behavior: Case study [World Conference on Information Technology]. *Procedia Computer Science*, *3*, 57–63. https://doi.org/https://doi.org/10.1016/j.procs.2010.12.011 (cit. on p. 43)

Kidd, S., Gelders, B., Bailey-Athias, D., et al. (2017). *Exclusion by design: An assessment of the effectiveness of the proxy means test poverty targeting mechanism* (tech. rep.). International Labour Organization. (Cit. on p. 20).

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. (1983). Optimization by simulated annealing. *Science*, *220*, 671–680 (cit. on p. 53).

Klievink, A. (2021). *Hollen én stilstaan: Hoe data en digitalisering de overheid veranderen* (Doctoral dissertation). Leiden University. (Cit. on p. 22).

Knippenberg, E., Jensen, N., & Constas, M. A. (2017). Resilience, shocks and the dynamics of well-being evidence from malawi∗ (cit. on p. 9).

Ko, S., Kang, S., Kang, H., & Lee, M. J. (2018). An exploration of foreign tourists' perceptions of korean food tour: A factor-cluster segmentation approach. *Asia Pacific Journal of Tourism Research*, *23*(8), 833–846 (cit. on p. 48).

Kotler, P., & Armstrong, G. (1999). *Principles of marketing*. Prentice Hall. https://books.google.nl/books?id=AxKVQgAACAAJ. (Cit. on p. 41)

Lambrecht, A., & Tucker, C. E. (2020). Apparent algorithmic discrimination and real-time algorithmic learning in digital search advertising. *Available at SSRN 3570076* (cit. on p. 22).

Leavy, S., O'Sullivan, B., & Siapera, E. (2020). Data, power and bias in artificial intelligence. *arXiv preprint arXiv:2008.07341* (cit. on p. 22).

Lefait, G., & Kechadi, T. (2010). Customer segmentation architecture based on clustering techniques. *2010 Fourth International Conference on Digital Society*, 243–248 (cit. on p. 44).

Lie, J. H. S. (2017). From humanitarian action to development aid in northern uganda and the formation of a humanitarian-development nexus. *Development in Practice*, *27*(2), 196–207. https://doi.org/10.1080/09614524.2017.1275528 (cit. on p. 18)

Lilien, G. L., & Kotler, P. (1983). *Marketing decision making: A model-building approach*. Harpercollins College Division. (Cit. on p. 41).

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, *1*(14), 281–297 (cit. on p. 52).

Maghawry, A., Al-qassed, A., Awad, M., & Kholief, M. (2021). Automated market analysis by rfmx encoding based customer segmentation using initial centroid selection optimized k-means clustering algorithm. *IJCI. International Journal of Computers and Information*, *8*(2), 26–31 (cit. on p. 43).

Makgosa, R., & Sangodoyin, O. (2018). Retail market segmentation: The use of consumer decision-making styles, overall satisfaction and demographics. *The International Review of Retail, Distribution and Consumer Research*, *28*(1), 64–91 (cit. on p. 42).

McBride, L., & Nichols, A. (2016). Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning. *The World Bank Economic Review*. https://doi.org/10.1093/wber/lhw056 (cit. on p. 3)

Meijer, V. (2021). From everything on paper to everything digital. (Cit. on p. 28).

Menéndez, J., & Barrena, I. (2021). *Menéndez, J., & Barrena, I. (2021, November). FINAL EVALUATION COVID-19 FOOD ASSISTANCE PROGRAM CAS ISLANDS* (tech. rep.). Egin International Consulting. (Cit. on pp. 4, 5, 21, 28, 33).

MercyCorps. (2007). *Guide to Cash-for-Work Programming.* (tech. rep.). MercyCorps. (Cit. on p. 5).

Miglautsch, J. (2000). Thoughts on rfm scoring. *The Journal of Database Marketing*, *8*, 67–72. https://doi.org/10.1057/palgrave.jdm.3240019 (cit. on p. 43)

Ministry of Economic Affairs, T., Tourism, & Telecommunications. (2020). *Country Sint Maarten Economic Indicators Half year developments 2020* (tech. rep.). Department of Statistics Sint Maarten. (Cit. on p. 26).

Monaghan, A., & Lycett, M. (2013). Big data and humanitarian supply networks: Can Big Data give voice to the voiceless? *2013 IEEE Global Humanitarian Technology Conference (GHTC)*, 432–437. https://doi.org/10.1109/GHTC.2013.6713725 (cit. on p. 9)

Monalisa, S., & Kurnia, F. (2019). Analysis of dbscan and k-means algorithm for evaluating outlier on rfm model of customer behaviour. *Telkomnika*, *17*(1), 110–117 (cit. on p. 44).

Moreno, R. (2017). *Year in Review 2017* (tech. rep.). UNHCR Innovation Service. (Cit. on pp. 3, 9).

Mulvenna, M., Buchner, A., Anand, S., & Hughes, J. (1999). Discovering & deploying marketing knowledge through web usage mining. *UNICOM 1999* (cit. on p. 43).

Myers, J. H., & Tauber, E. (1977). Market structure. (Cit. on p. 41).

Nature Foundation Sint Maarten. (2016). Seagrass. (Cit. on p. 26).

Netherlands Red Cross. (2017). *First public report about the national campaign "The Netherlands helps St. Maarten"* (tech. rep.). Red Cross Netherlands. (Cit. on p. 27).

Netherlands Red Cross. (2018). *Second public report about the national campaign 'The Netherlands helps St. Maarten'* (tech. rep.). Netherlands Red Cross. (Cit. on pp. 8, 27).

Netherlands Red Cross. (2019). *Vierde terugkoppeling Nationale Actie "Nederland helpt Sint-Maarten"* (tech. rep.). Red Cross Netherlands. (Cit. on p. 27).

Netherlands Red Cross. (2020a). *Cash feasibility analysis COVID19 - Sint Maarten* (tech. rep.). Netherlands Red Cross. (Cit. on p. 28).

Netherlands Red Cross. (2020b). *Vijfde terugkoppeling Nationale Actie "Nederland helpt Sint-Maarten"* (tech. rep.). Netherlands Red Cross. (Cit. on p. 27).

Netherlands Red Cross. (2021a). *Nationaal Actieplan COVID-19 RAMP OP RAMP* (tech. rep.). Red Cross Netherlands. (Cit. on p. 27).

Netherlands Red Cross. (2021b). *STRATEGIE 2021 - 2025 VERSTERKEN VAN DE MENSELIJKE BASIS* (tech. rep.). Red Cross Netherlands. The Hague. (Cit. on p. 30).

Netherlands Red Cross. (2021c). *Voedselpakket vs. voucher: Waarom je mensen soms beter cash dan een hulppakket kunt geven.* https://www.rodekruis.nl/nieuwsbericht/voedselpakket-vs-voucher-waarom-je-mensen-soms-beter-cash-dan-een-hulppakket-kunt-geven/. (Cit. on p. i)

Nightingale, K. (2012). *Building the future of humanitarian aid: Local capacity and partnerships in emergency assistance* (tech. rep.). Christian Aid. (Cit. on p. 1).

Norton, R., MacClune, K., & Szönyi, M. (2020). *When the unprecedented becomes precedented: Learning from Cyclones Idai and Kenneth.* (tech. rep.). ISET International and the Zurich Flood Resilience Alliance. Boulder, CO. (Cit. on p. 1).

Obrecht, A. (2017). *Using Evidence to Allocate Humanitarian Resources: Challenges and Opportunities* (tech. rep.). ALNAP/ODI. London. (Cit. on pp. 3, 18, 73).

Ojiambo, S., & Chamaa, C. (2021). *Getting CVA right in the Humanitarian Response Plan: Blind Spots and Considerations.* (tech. rep.). Cash Learning Partnership. (Cit. on p. 4).

O'Neil, C., & Schutt, R. (2013). *Doing data science: Straight talk from the frontline.* " O'Reilly Media, Inc.". (Cit. on pp. 10, 11).

Pallant, J. (2020). *Spss survival manual: A step by step guide to data analysis using ibm spss.* Routledge. (Cit. on p. 50).

Peachey, K. (2021). A short history of cash and voucher assistance – 6 key lessons and observations. (Cit. on p. 2).

Preud'Homme, G., Duarte, K., Dalleau, K., Lacomblez, C., Bresso, E., Smaïl-Tabbone, M., Couceiro, M., Devignes, M.-D., Kobayashi, M., Huttin, O., et al. (2021). Head-to-head comparison of clustering methods for heterogeneous data: A simulation-driven benchmark. *Scientific reports*, *11*(1), 1–14 (cit. on p. 52).

Rhys, H. (2020). *Machine Learning with R, the tidyverse, and mlr* (Vol. 1). Manning Publications. (Cit. on p. 47).

Ritchie, H., & Roser, M. (2020). *CO and greenhouse gas emissions.* (tech. rep.). Our world in data. (Cit. on p. 1).

Roger, R. (2007). *Does Foreign Aid Really Work?* New York: Oxford University Press. (Cit. on p. 22).

Royal Netherlands Institute for Sea Research. (2020). Caribbean islands face loss of protection and biodiversity as seagrass loses terrain. (Cit. on p. 26).

Rundle-Thiele, S., Kubacki, K., Tkaczynski, A., & Parkinson, J. (2015). Using two-step cluster analysis to identify homogeneous physical activity groups. *Marketing Intelligence & Planning* (cit. on p. 41).

Sebastian, A. R., Shivakumaran, S., Silwal, A. R., Newhouse, D. L., Walker, T. F., & Yoshida, N. (2018). A proxy means test for sri lanka. *World Bank Policy Research Working Paper*, (8605) (cit. on p. 20).

Seltzer, W. (2006). The dark side of numbers: Updated. In R. Mackensen (Ed.), *Bevölkerungsforschung und politik in deutschland im 20. jahrhundert* (pp. 119–136). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90427-6_7. (Cit. on p. 22)

Smith, S. (1989). Tourism analysis: A handbook. 322 s. (Cit. on p. 48).

Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of marketing*, *21*(1), 3–8 (cit. on p. 41).

Soualiga News Today. (2017). Red cross food voucher st. maarten. https://www.soualiganewsday.com/index.php?option=com_k2&view=item&id=16531:red-cross-launches-food-voucher-program-to-support-12,025-hurricane-irma-survivors&Itemid=450. (Cit. on p. 27)

Sphere Project. (2011). *Sphere handbook: Humanitarian charter and minimum standards in disaster response.* https://www.refworld.org/docid/4ed8ae592.html. (Cit. on p. 17)

Sterling, T. (2017). Netherlands PM: Death toll from Irma on Dutch Saint Martin rises to four. (Cit. on p. 26).

StMaartenNews.com. (2020). St. Maarten informal economy grounded to a halt. (Cit. on p. 25).

Stodden, V. (2020). The data science life cycle: A disciplined approach to advancing data science as a science. *Communications of the ACM*, *63*(7), 58–66 (cit. on p. 10).

The Inter-Agency Standing Committee. (2021). *The Grand Bargain 2.0 Endorsed framework and annexes.* (tech. rep.). (Cit. on p. 4).

The World Bank Group. (2016). *Strategic Note: Cash Transfers in Humanitarian Contexts* (tech. rep.). The World Bank Group. Washington, DC. (Cit. on pp. 2, 21).

Tibshirani, R., & Walther, G. (2005). Cluster validation by prediction strength. *Journal of Computational and Graphical Statistics*, *14*(3), 511–528 (cit. on p. 53).

UN OCHA. (2019). SINT MAARTEN (SXM) Common Operational Dataset Administrative Boundaries. (Cit. on p. 30).

UN OCHA. (2021). *Global Humanitarian Overview 2021. Geneva, Switzerland: The United Nations Office for the Coordination of Humanitarian Affairs (OCHA).* (tech. rep.). The United Nations Office for the Coordination of Humanitarian Affairs. Geneva, Switzerland. (Cit. on p. 1).

UNFPA. (2020). *Humanitarian Cash and Voucher Assistance (CVA) Tip Sheet: CVA Overview* (tech. rep.). https://gbvaor.net/sites/default/files/2020-05/UNFPA%20CVA%20Overview%20Tip%20Sheet_final.pdf. (Cit. on p. 21)

UNICEF. (2020). *Situation Analysis on Children and Adolescents on Sint Maarten 2020* (tech. rep.). United Nations Children's Fund. https://www.unicef.nl/files/Situation%20Analysis_St%20Maarten_2020_Full%20Report_EN.pdf. (Cit. on p. 25)

Van den Homberg, M., Nobre, G. G., Veldkamp, T., Bolton, T., Bischiniotis, K., Davenport, F., Ambani, M. K., Abdillahi, H. S., & Aerts, J. (2019). Forecast based financing for food security. *Geophysical Research Abstracts*, *21* (cit. on pp. 3, 18).

Van der Mee, T. (2020). Na orkaan Irma stort corona Sint Maarten in nog grotere crisis. (Cit. on p. 8).

Verme, P., & Gigliarano, C. (2019). Optimal targeting under budget constraints in a humanitarian context. *World Development*, *119*, 224–233 (cit. on p. 3).

Veynberg, R., Timofeev, A., Popov, A., & Bortsova, D. (2018). Data driven marketing as a new approach to business development and sales methods. *Espacios*, *39*(12), 3 (cit. on p. 43).

Vogel, B., Tschunkert, K., & Schläpfer, I. (2022). The social meaning of money: Multidimensional implications of humanitarian cash and voucher assistance. *Disasters*, *46*(2), 348–370. https://doi.org/https://doi.org/10.1111/disa.12478 (cit. on p. 1)

Voncken, H. (2020). Sint-Maarten op slot tegen coronavirus: bedrijven en scholen ook dicht. (Cit. on p. 26).

Vos, N. J. D. (2021). Kmodes categorical clustering library. (Cit. on p. 54).

Wedel, M., & Kamakura, W. (2012). *Market segmentation: Conceptual and methodological foundations* (Vol. 8). https://doi.org/10.1007/978-1-4615-4651-1. (Cit. on pp. 41, 42, 44)

Weinstein, A. (1987). *Market segmentation: Using demographics, psychographics, and other segmentation techniques to uncover and exploit new markets*. Irwin Professional Publishing. (Cit. on p. 41).

Weinstein, A. (2006). A strategic framework for defining and segmenting markets. *Journal of Strategic Marketing*, *14*(2), 115–127 (cit. on p. 42).

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian journal of paramedicine*, *8*(3) (cit. on p. 50).

World Bank. (2021). Sint Maarten Overview: Development news, research, data. (Cit. on p. 26).

World Humanitarian Summit. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need.* (tech. rep.). High-Level Panel on Humanitarian Financing. Istanbul. (Cit. on pp. 1, 2).

WorldData.info. (2019). Development and importance of tourism for Sint Maarten. (Cit. on pp. 25, 26).

WorldData.info. (2022). Sint Maarten: country data and statistics. https://www.worlddata.info/america/sint-maarten/index.php. (Cit. on p. 25)

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., & Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nature communications*, *11*(1), 1–11 (cit. on pp. 3, 18).

Yousaf, S., & Huaibin, L. (2013). Profiling consumer behavior in the context of involvement level and demographic factors: Evidence of within-country differences from a developing economy. *Journal of Global Marketing*, *26*(1), 1–17 (cit. on p. 42).

A
Appendix: Data Flow Diagram

Figure A.1 shows the processes of the used data in this research.

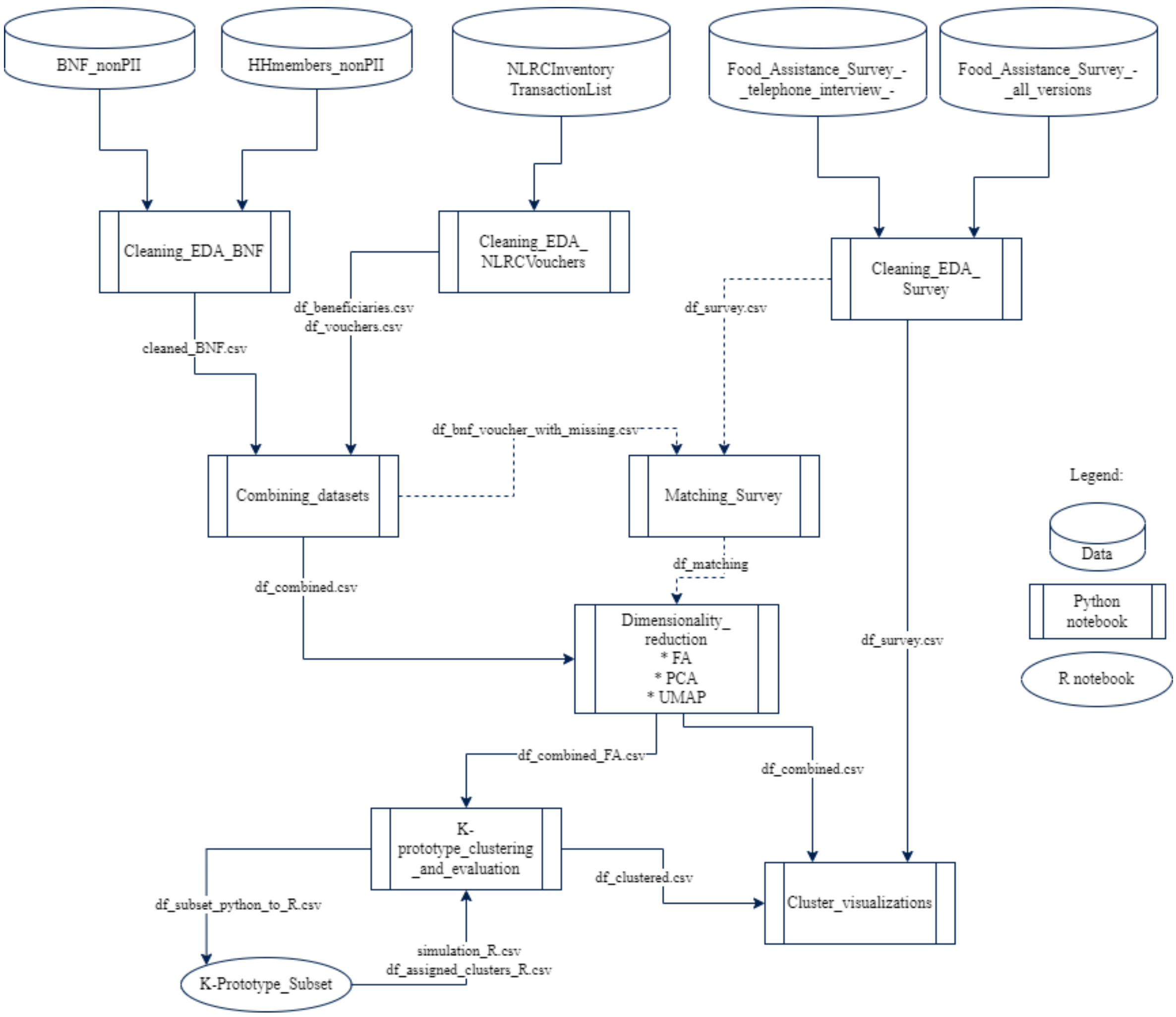


Figure A.1: Data Flow Diagram

B

Appendix: Heatmap

The heatmap in figure B.1 provides correlations between pairs of selected variables in the subset. It has been used in the factor analysis for choosing the highest correlating variables.

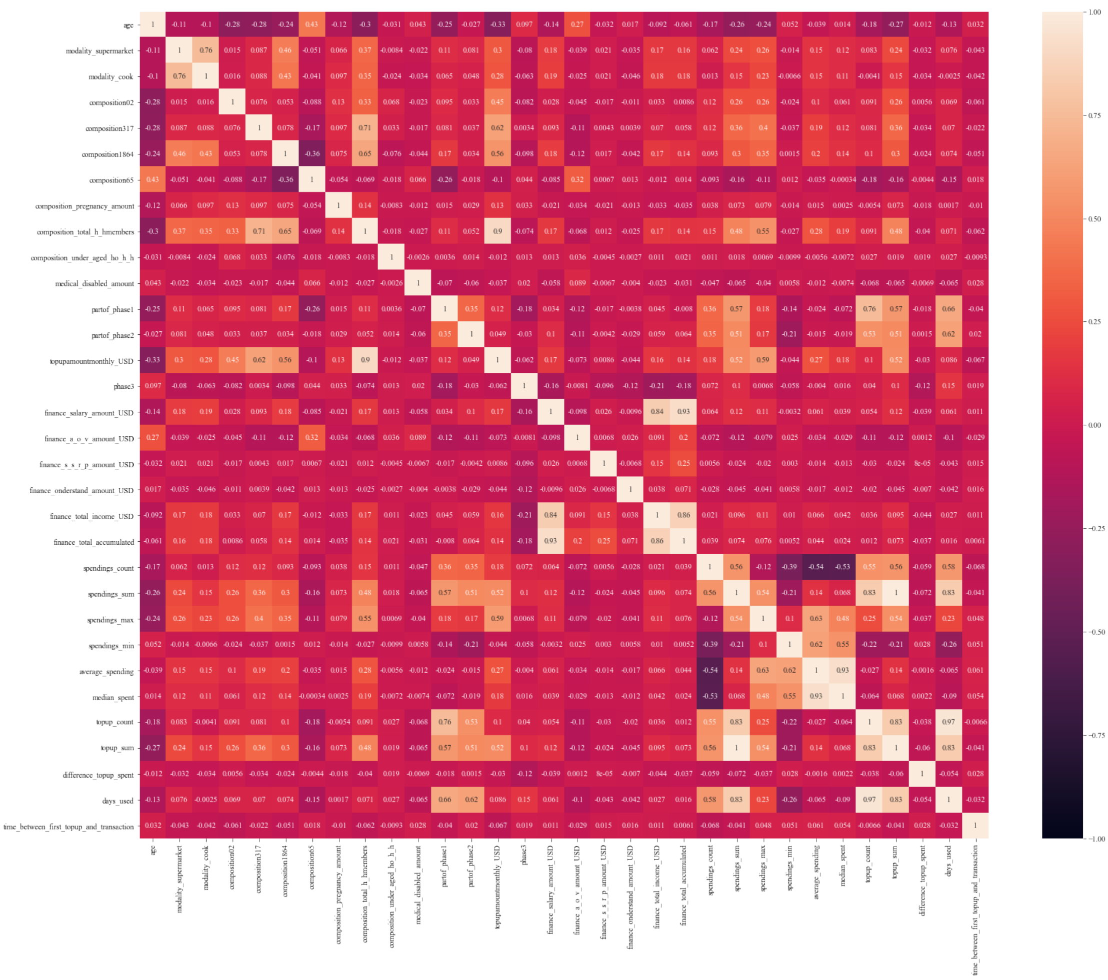
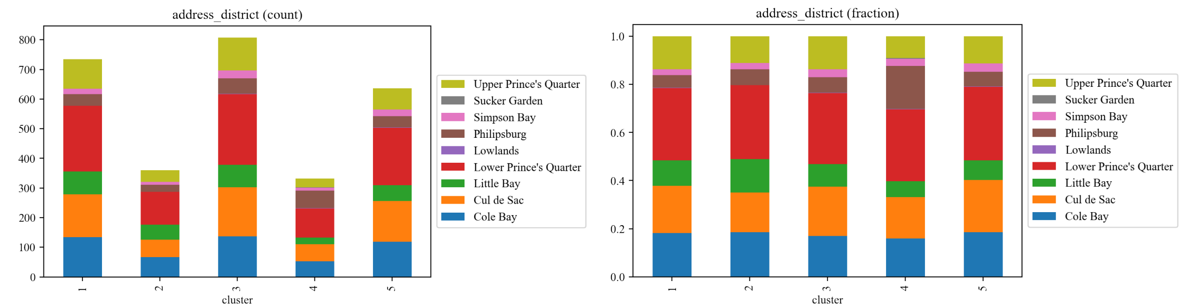


Figure B.1: Heatmap showing all correlations between pairs of variables

C

Appendix: Cluster Results

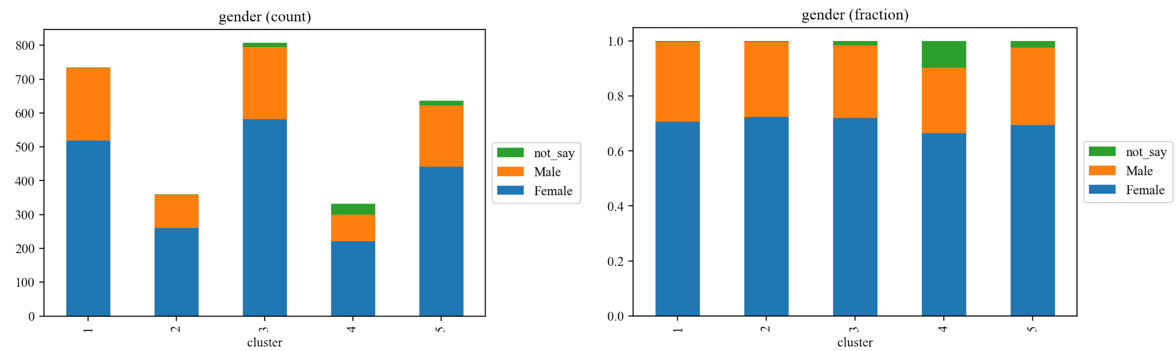
This appendix presents different graphs which are used in the cluster analysis. Figures C.1 to C.6b are used for analysing categorical variables. This has been done by comparison of frequencies in absolute numbers and in fractions. Figures C.7 until C.11 are used for inspecting absolute values of numerical variables using radarplots.



(a) Absolute

(b) Fractions

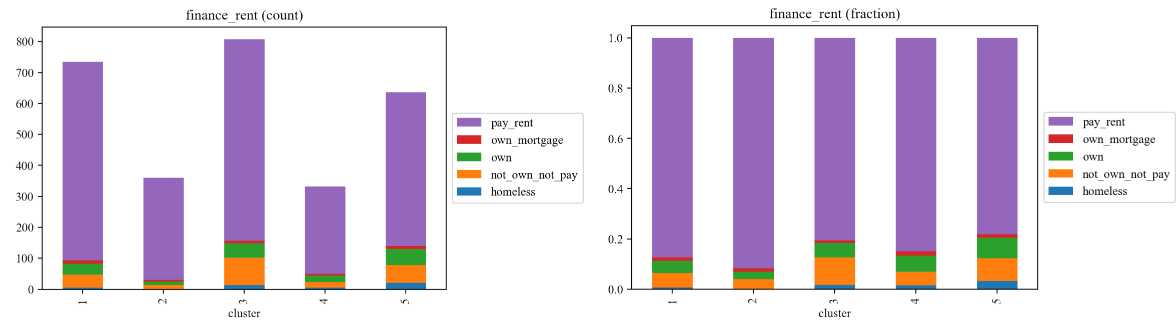
Figure C.1: Distribution of districts



(a) Absolute

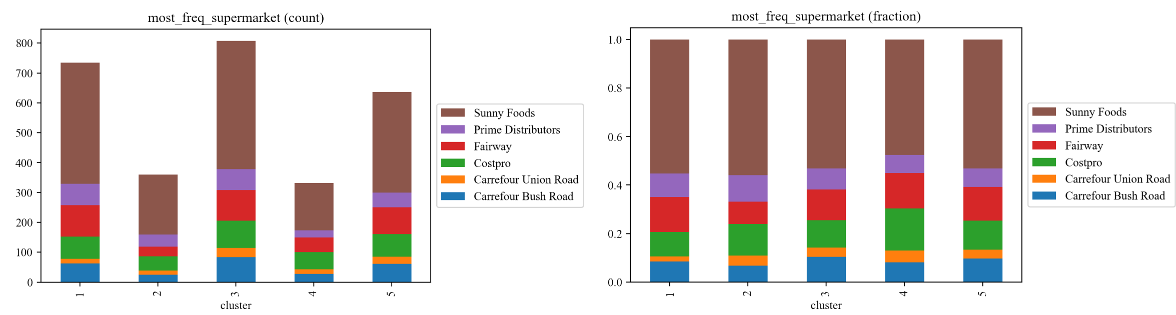
(b) Fractions

Figure C.2: Distribution of gender



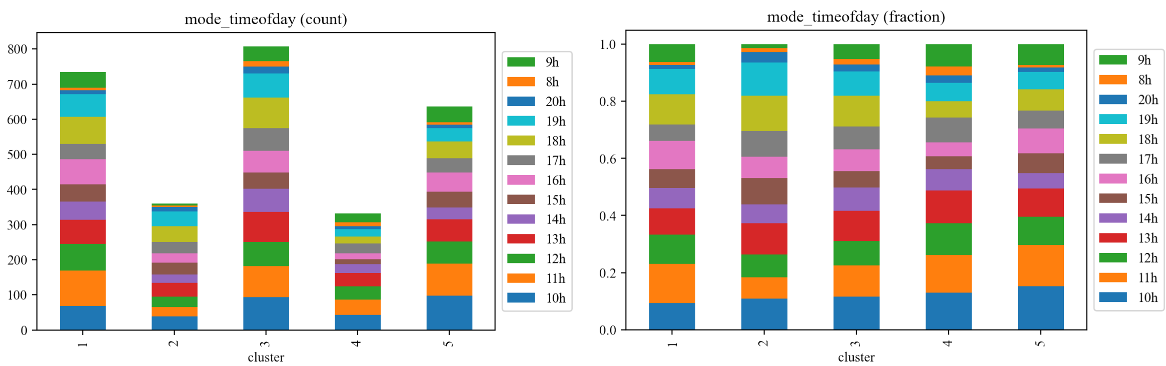
(a) Absolute **(b)** Fractions

Figure C.3: Distribution of living situation



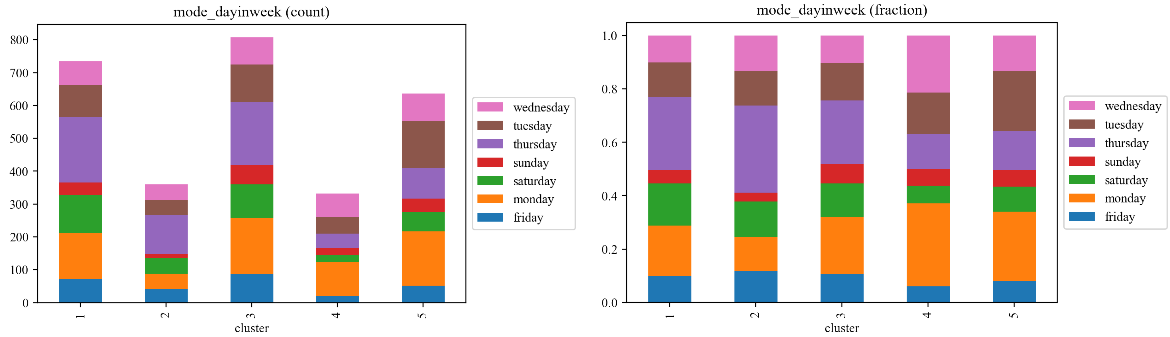
(a) Absolute **(b)** Fractions

Figure C.4: Distribution of most frequently visited supermarket



(a) Absolute **(b)** Fractions

Figure C.5: Distribution of most frequently visited time of day



(a) Absolute

(b) Fractions

Figure C.6: Distribution of most frequently visited day in week

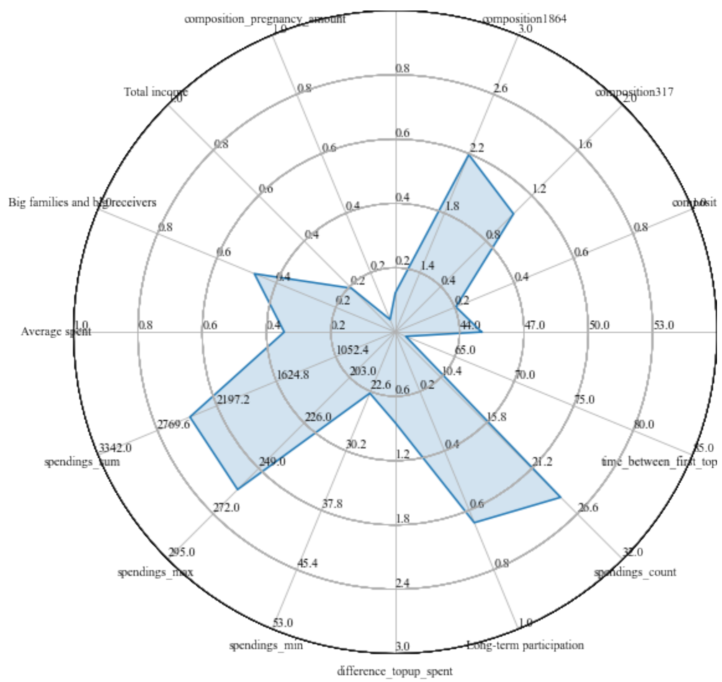


Figure C.7: Radarplot on cluster 1 using absolute values

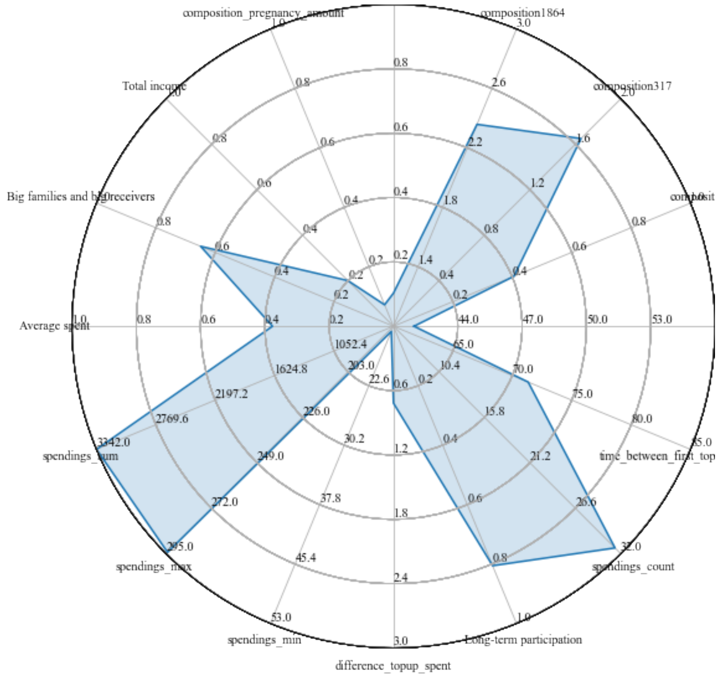


Figure C.8: Radarplot on cluster 2 using absolute values

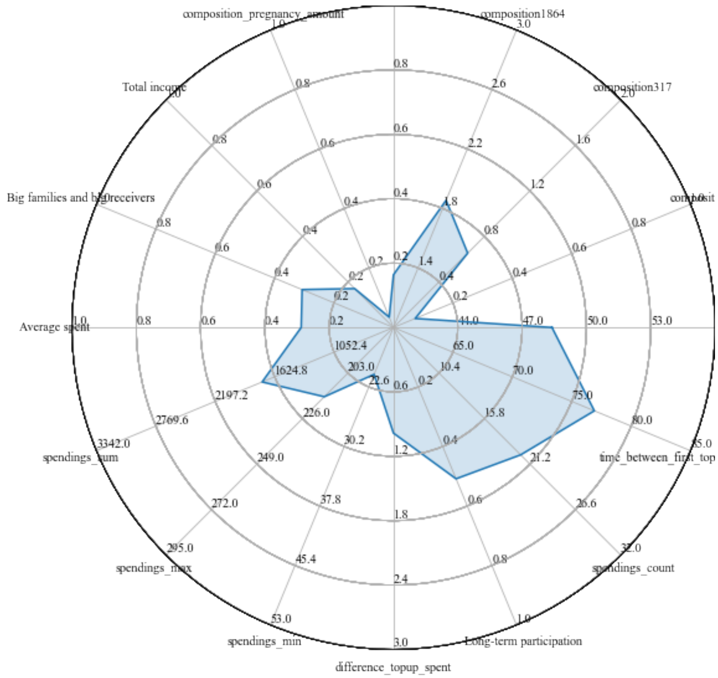


Figure C.9: Radarplot on cluster 3 using absolute values

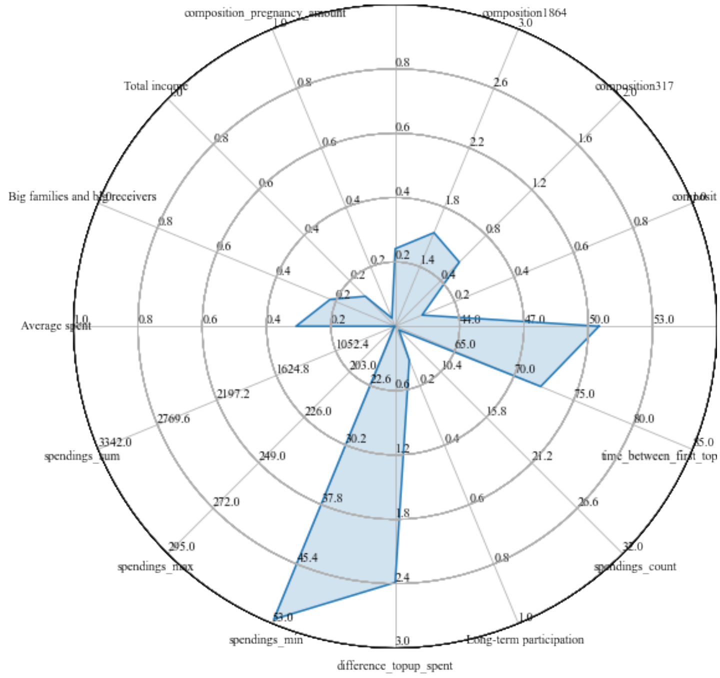


Figure C.10: Radarplot on cluster 4 using absolute values

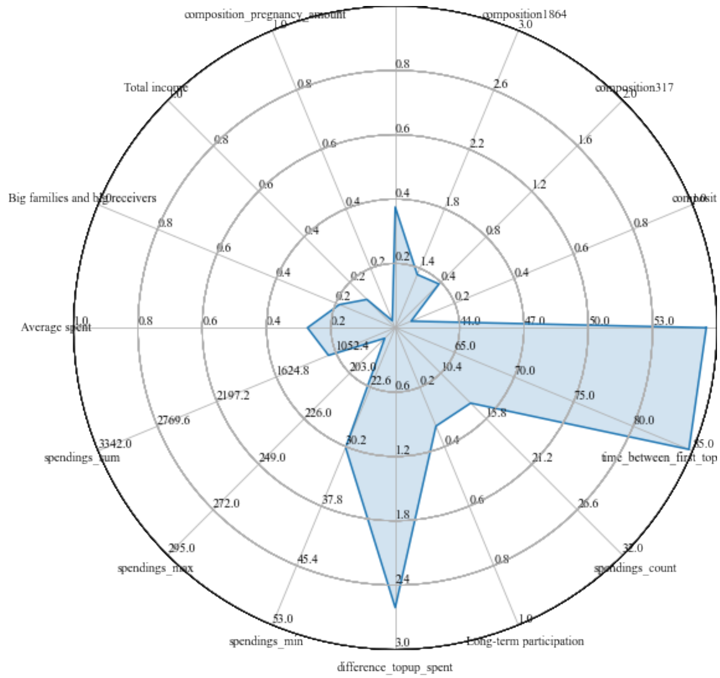


Figure C.11: Radarplot on cluster 5 using absolute values

D

Appendix: Available Data

This is an overview of the available datasets of the evoucher project in Sint Maarten. It consists of participant demographics, voucher transaction data, supermarket receipt data and of two surveys after the project has been finished (one was filled in individually and one is conducted via an interview by telephone). All of these data do not only consist of people who received evouchers, but also of people who received hot meals or food boxes, or a combination of all the aforementioned. This appendix is used for understanding the available datasets and for exploring values, units of measurement and to see how much data is available. When useful, some explanation has been given of every column of the dataset.

D.1. Food Assistance Survey (Post Distribution Monitoring)

This survey is a raw dataset and will be cleaned. There is a jump from one part of the survey to another part. This happens when question 5 is answered with 'No' (so participant of survey is not a participant of food assistance) which skip questions up to question 98. If question 5 is answered with 'Yes', then the survey will finish at question 97. The dimensions are: 1677 rows (around 749 relevant, only NLRC evoucher) and 140 columns (around 123 relevant).

Table D.1: Description of survey data (collected digitally)

Question number	Column name	Explanation	Possible values	Example value 1	Example value 2	Example value 3
1	start	Start time of taking survey	Year-month-day-time	2020-12-09T16:10:58.570-04:00	2020-12-11T18:32:08.006-04:00	2020-12-09T16:44:38.236-04:00
2	end	End time of taking survey	Year-month-day-time	2020-12-09T16:27:45.114-04:00	2020-12-11T18:43:27.194-04:00	2020-12-09T16:48:26.623-04:00
3	today	Date	Year-month-day	08/12/2020	10/12/2020	08/12/2020
4	Language can be changed on top	-	(blank)	0	0	0
5	Did you receive food assistance in the year 2020 from the Red Cross, K1, Freegan, Captains Rib Shack or	If answered no, this would skip most of the questionnaire and go to the last section. See question	Yes No	No	Yes	No
6	The Food Collective would like to get your opinion about the food assistance that you are receiving. We will use this information to get an insight in your satisfaction with the programme and your feedback will help us improve our programme. The information that you share is completely anonymous and only aggregated data will be shared with external actors. The survey will take approximately 10 minutes. Would like to continue with the questionnaire? By continuing you give your consent that we can use your information to generate aggregated data to be used for internal and external reporting.	Consent, speaks for itself. Most people consented luckily.	Yes No (blank)	0	Yes	0
7	What is your age?		[2…19111955]	0	52	0
8	What is your nationality?		Dutch Dominican Other, specify: Hatian Dominica Venezualian Colombian (blank)	0	Dominican	0
			Jamaican Guyanese			

9	Specify your nationality		St.lucian Jamaica Indian St lucian [and more]	0	0	0
10	How many people in the household are aged between 0 and 2		Most between [0…5]	0	0	0
11	How many people in the household are aged between 3 and 17		Most between [0…5]	0	2	0
12	How many people in the household are aged between 18 and 65		Most between [0…5]	0	2	0
13	How many people in the household are aged 65+		Most between [0…5]	0	0	0
14	total_hh_members		[1…6]	0	4	0
15	The total of members (including you) is ${total_hh_members}. Is that right?		OK (blanks)	0	OK	0
16	Did you receive at any time during the year:		Various	0	Other, specify:	0
17	Did you receive at any time during the year:/SSRP support		[0] = No [1] = Yes	0	0	0
18	Did you receive at any time during the year:/Social Assistance / onderstand		[0] = No [1] = Yes	0	0	0
19	Did you receive at any time during the year:/Pension (AOV)		[0] = No [1] = Yes	0	0	0
20	Did you receive at any time during the year:/Financial support from friends or family from abroad		[0] = No [1] = Yes	0	0	0
21	Did you receive at any time during the year:/Financial support from friends or family locally		[0] = No [1] = Yes	0	0	0
22	Did you receive at any time during the year:/Public Health insurance		[0] = No [1] = Yes	0	0	0
23	Did you receive at any time during the year:/Other, specify:		[0] = No [1] = Yes	0	1	0
24	Specify your assistance		Various	0	Cruz Roja	0
25	Which food assistance did you receive during the last 6 months?		One or combination of: Food box Freegan, Food box K1, Hot Meal Captains Rib Shack, Hot Meal COME center, Evoucher K1, Evoucher NLRC	0	Evoucher NLRC	0
26	Which food assistance did you receive during the last 6 months?/Food box Freegan		[0] = No [1] = Yes	0	0	0
27	Which food assistance did you receive during the last 6 months?/Food box K1		[0] = No [1] = Yes	0	0	0

28	Which food assistance did you receive during the last 6 months?/Hot Meal Captains Rib		[0] = No [1] = Yes	0	0	0
29	Which food assistance did you receive during the last 6 months?/Hot Meal COME center		[0] = No [1] = Yes	0	0	0
30	Which food assistance did you receive during the last 6 months?/Evoucher NLRC	Most relevant here since this is the evoucher from the NLRC	[0] = No [1] = Yes	0	1	0
31	Which food assistance did you receive during the last 6 months?/Evoucher K1		[0] = No [1] = Yes	0	0	0
32	During registration we've asked you for your preferred type of food assistance, did you receive food assistance of your preference?		Yes, No	0	Yes	0
33	What did you prefer?		0, Evoucher, Food parcels	0	0	0
34	Do you know on what basis you were selected for the food assistance programme?		Yes, No	0	Yes	0
35	How do you rate the communication before receiving assistance (time of distribution, location, modality)?		good very good bad very bad	0	very good	0
36	What could be improved?	Possibility for writing recommendations		0	0	0
37	How do you rate the quality of the distribution (waiting time, information provided during distribution, organization)		0 good very good bad very bad	0	very good	0
38	What could be improved?	Possibility for writing recommendations		0	0	0
39	Did you feel safe during the distribution?		0 very safe somewhat safe not safe	0	very safe	0
40	What could be improved?	Possibility for writing recommendations		0	0	0
41	Did you feel that the staff was friendly and helpful during the distribution?		0 very friendly/helpful somewhat friendly/helpful not friendly/helpful	0	very friendly/helpful	0
42	What could be improved?	Possibility for writing recommendations		0	0	0
43	Did anyone ever asked you to pay anything in order to receive food assistance?		0 No Yes	0	No	0
			0			

44	Was the money provided on the evoucher enough for your household to buy food and hygiene items?		very sufficient somewhat sufficient not sufficient	0	somewhat sufficient	0
45	Did you ever experience any difficulties with using the evouchers in the supermarket?		0 No Yes	0	No	0
46	What difficulties did you experience?	Personal clarifications		0	0	0
47	What difficulties did you experience?/Card didn't work		[0] = No [1] = Yes	0	0	0
48	What difficulties did you experience?/Lost card		[0] = No [1] = Yes	0	0	0
49	What difficulties did you experience?/Forgot pin code		[0] = No [1] = Yes	0	0	0
50	What difficulties did you experience?/Forgot 5 letter code		[0] = No [1] = Yes	0	0	0
51	What difficulties did you experience?/Took long time to do the transaction		[0] = No [1] = Yes	0	0	0
52	What difficulties did you experience?/Return to supermarket because of wrong transaction		[0] = No [1] = Yes	0	0	0
53	What difficulties did you experience?/Supermarket Staff was not helpful		[0] = No [1] = Yes	0	0	0
54	What difficulties did you experience?/Other (specify)		[0] = No [1] = Yes	0	0	0
55	Specify your difficulty:	Personal clarifications	Various	0	0	0
56	Did you experience any increase in prices in the supermarkets since you started using the voucher?		0 No Yes	0	Yes	0
57	How much did prices increase		0 small increace some increase large increase	0	some increase	0
58	What items did you mostly buy with the evoucher?		One or combination of: Meat Beverages Vegetables Canned food Other (specify) Cleaning/hygiene items Dry food (rice, legumes)	0	Other (specify)	0
59	What items did you mostly buy with the evoucher?/Vegetables		[0] = No [1] = Yes	0	0	0
60	What items did you mostly buy with the evoucher?/Beverages		[0] = No [1] = Yes	0	0	0
61	What items did you mostly buy with the evoucher?/Dry food (rice		[0] = No	0	0	0

61	the evoucher?/Dry food (rice, legumes)		[1] = Yes	0	0	0
62	What items did you mostly buy with the evoucher?/Canned food		[0] = No [1] = Yes	0	0	0
63	What items did you mostly buy with the evoucher?/Meat		[0] = No [1] = Yes	0	0	0
64	What items did you mostly buy with the evoucher?/Cleaning/hygiene items		[0] = No [1] = Yes	0	0	0
65	What items did you mostly buy with the evoucher?/Other (specify)		[0] = No [1] = Yes	0	1	0
66	Specify the items you mostly bought with the evoucher		Various	0	Alimentos de consumo diario preparar la comidaara pr	0
67	Do you have any suggestions on how to improve the evoucher assistance?		No, Yes	0	No	0
68	Please give your suggestions:		Various	0	0	0
69	How would you rate the quality of the meals?	Not relevant		0	0	0
70	What would you like to change	Not relevant		0	0	0
71	Were the portions sufficient for you?	Not relevant		0	0	0
72	Were the meals enough to meet the food needs in your household?	Not relevant		0	0	0
73	Do you have any suggestions on how to improve the daily hot meals?	Not relevant		0	0	0
74	Please give your suggestions:	Not relevant		0	0	0
75	How would you rate the quality of the food boxes?	Not relevant		0	0	0
76	What would you like to change	Not relevant		0	0	0
77	Was there enough food provided to meet the food needs in your household?	Not relevant		0	0	0
78	Would you like to change the food items that are provided in the food boxes	Not relevant		0	0	0
79	What would you like to change	Not relevant		0	0	0
80	Do you have any suggestions on how to improve the foodboxes?	Not relevant		0	0	0
81	Please give your suggestions:	Not relevant		0	0	0
82	How would you rate the quality of the hygiene items?	Not relevant		0	0	0

83	Were there enough hygiene items to meet the needs in your household?	Not relevant		0	0	0
84	Would you like to change the hygiene items that are provided?	Not relevant		0	0	0
85	What would you like to change?	Not relevant		0	0	0
86	Do you know how to reach the Red Cross, K1, Freegan, Captains Rib Shack or COME Center when you		Yes No	0	Yes	0
87	Did you ever contact the Red Cross, K1, Freegan, Captains Rib Shack or COME Center?		Yes No	0	Yes	0
88	How did you contact us?		Call by Phone Other (specify) Send a WhatsApp Sent an email Visit the office	0	Send a WhatsApp	0
89	How did you contact us?/Call by Phone		[0] = No [1] = Yes	0	0	0
90	How did you contact us?/Send a WhatsApp		[0] = No [1] = Yes	0	1	0
91	How did you contact us?/Visit the office		[0] = No [1] = Yes	0	0	0
92	How did you contact us?/Sent an email		[0] = No [1] = Yes	0	0	0
93	How did you contact us?/Other (specify)		[0] = No [1] = Yes	0	0	0
94	Specify how you contacted us:		Various	0	0	0
95	When you contacted Red Cross, K1, Freegan, Captains Rib Shack or COME Center, was the problem		Yes No	0	Yes	0
96	Please specify your problem:		Various	0	0	0
97	Please click validate to finalise your survey. Thank you for filling out the survey and providing us with valuable feedback.		0	0	0	0

98	The food collective deems it essential to obtain information from people not included in our food assistance program to provide an insight on the direction of the Food Assistance Project. The information that you share is completely anonymous and only aggregated data will be shared with external actors. Please take into consideration that providing your insight on your household circumstance does not lead to automatic food assistance nor being assigned to a waiting list. Would like to continue with the questionnaire? By continuing you give your consent that we can use your information to generate aggregated data to be used for internal and external reporting.	This question is only answered to people who did not receive food assistance in the year 2020 from the Red Cross, K1, Freegan, Captains Rib Shack or COME Center.	Yes No	Yes	0	Yes
99	What is your age?		18 to 75	52	0	63
100	What is your nationality?		Dutch Dominican 0 Other, specify: Venezualian Hatian Dominica	Dominican	0	Other, specify:
101	Specify your nationality		0 Jamaican Haitian Jamaica Guyananese St. Kitts Guyanese Guyanaese Grenadian French Kittitian	0	0	Jamaican
102	Did you or someone in your household have work before April 2020 and lost the job because of the		Yes No	Yes	0	Yes
			Other (specify)	Buying less		

103	Did you have any financial problems over the last 6 months?		Behind on rent payments Buying less or cheaper food Buying less or no fuel Buying less or no clothes	Buying less or cheaper food Buying less or no clothes Buying less or no fuel	0	Buying less or cheaper food
104	Did you have any financial problems over the last 6 months?/Behind on rent payments		[0] = No [1] = Yes	0	0	0
105	Did you have any financial problems over the last 6 months?/Buying less or cheaper food		[0] = No [1] = Yes	1	0	1
106	Did you have any financial problems over the last 6 months?/Buying less or no clothes		[0] = No [1] = Yes	1	0	0
107	Did you have any financial problems over the last 6 months?/Buying less or no fuel		[0] = No [1] = Yes	1	0	0
108	Did you have any financial problems over the last 6 months?/Other (specify)		[0] = No [1] = Yes	0	0	0
109	Please specify your financial problem:		Various	0	0	0
110	If you would receive food assistance which modality would you have chosen?		0 eVoucher Food Box Other (specify) Hot Meal	eVoucher	0	eVoucher
111	Please specify what food-assistance you would've liked:	Not relevant		0	0	0
112	Is there a member in your household with specific dietary needs		Yes No	No	0	Yes
113	Is there a member in your household with chronic disability and is not mobile enough to cook or go to		Yes No	No	0	No
114	Please take into consideration, that providing your insight on your household circumstance does not lead to automatic food assistance nor being assigned to a waiting list. Thank you for providing your insight. Click validate to finalise your survey.		0	0	0	0
115	Do you have any other suggestions on what we can do better?		Various	0	0	0
116	Thank you for your time		0	0	0	0
117	Did you receive at any time this year:	Empty		0	0	0
118	Did you receive at any time this year:/SSRP support	Empty		0	0	0
119	Did you receive at any time this year:/Social Assistance / onderstand	Empty		0	0	0

120	Did you receive at any time this year:/Pension (AOV)	Empty		0	0	0
121	Did you receive at any time this year:/Financial support from friends or family from abroad	Empty		0	0	0
122	Did you receive at any time this year:/Financial support from friends or family locally	Empty		0	0	0
123	Did you receive at any time this year:/Public Health insurance	Empty		0	0	0
124	Did you receive at any time this year:/Other, specify:	Empty		0	0	0
125	Specify your assistance	Empty		0	0	0
126	What difficulties did you experience?	Empty		0	0	0
127	What difficulties did you experience?/Card didn't work	Empty		0	0	0
128	What difficulties did you experience?/Lost card	Empty		0	0	0
129	What difficulties did you experience?/Forgot pin code	Empty		0	0	0
130	What difficulties did you experience?/Forgot 5 letter code	Empty		0	0	0
131	What difficulties did you experience?/Took long time to do the transaction	Empty		0	0	0
132	What difficulties did you experience?/Return to supermarket because of wrong transaction	Empty		0	0	0
133	What difficulties did you experience?/Supermarket Staff was not helpful	Empty		0	0	0
134	What difficulties did you experience?/Other (specify)	Empty		0	0	0
135	_id	Personal ID		37	1463	39
136	_uuid	User ID made by software for survey. Not relevant for research.		1289c1bd-fa65-4c23-b5ca-e85a75a0e6f7	adb67a6b-3499-4dcb-8b05-2c42bf3488d8	aa27363a-9b1c-42cb-bea9-8b0442acce33
137	_submission_time	Time		2020-12-09T20:27:47	2020-12-11T22:43:40	2020-12-09T20:48:30
138	_validation_status	Empty	0	0	0	0
139	_index	Number of entry	0 to 1680	1	1384	3

D.2. Food Assistance Survey Telephone Interview (Post Distribution Monitoring)

This survey is mostly the same as the Food Assistance Survey, with the difference that in the Food Assistance Survey people who did not receive assistance were also asked to participate in the survey. The telephone survey is only for people who are known to have been participating in this project while they did not fill in the other survey. The selection of people called for this survey was done from a representative sample. The dimensions of this dataset are: 106 columns (around 95 potentially relevant) and 73 rows (around 50 potentially relevant).

Table D.2: Description of survey data (collected via telephone)

Column name	Explanation	Possible values	Example 1	Example 2
start	Start time of taking survey?	Year-month-day-time	2020-12-08T11:58:17.756-04:00	2020-12-09T08:49:49.938-04:00
end	End time of taking survey?	Year-month-day-time	2020-12-08T12:07:00.739-04:00	2020-12-09T09:13:40.840-04:00
today	Date	Year-month-day	07/12/2020	08/12/2020
The Food Collective would like to get your opinion about the food assistance that you are receiving. We will use this information to get an insight in your satisfaction with the programme and your feedback will help us improve our programme. The information that you share is completely anonymous and only aggregated data will be shared with external actors. The survey will take approximately 10 minutes. Would like to continue with the questionnaire? By continuing you give your consent that we can use your information to generate aggregated data to be used for internal and external reporting.	Consent, speaks for itself. Most people consented luckily.	Yes No (blank)	Yes	Yes
What is your age?		Various	32	40
What is your nationality?		Dutch Dominican Other, specify: Hatian Dominica Venezualian Colombian (blank)	Hatian	Dominican
Specify your nationality		Jamaican Guyanese St.lucian Jamaica Indian St lucian [and more]	0	0
How many people in the household are aged between 0 and 2		Most between [0…5]	2	0
How many people in the household are aged between 3 and 17		Most between [0…5]	0	2
How many people in the household are aged between 18 and 65		Most between [0…5]	2	2
How many people in the household are aged 65+		Most between [0…5]	0	0
total_hh_members		[1…6]	4	4
The total of members (including you) is ${total_hh_members}. Is that right?		OK (blanks)	OK	OK

Did you receive at any time during the year:		Various	Financial support from friends or family locally	Other, specify:
Did you receive at any time during the year:/SSRP support		[0] = No [1] = Yes	0	0
Did you receive at any time during the year:/Social Assistance / onderstand		[0] = No [1] = Yes	0	0
Did you receive at any time during the year:/Pension (AOV)		[0] = No [1] = Yes	0	0
Did you receive at any time during the year:/Financial support from friends or family from abroad		[0] = No [1] = Yes	0	0
Did you receive at any time during the year:/Financial support from friends or family locally		[0] = No [1] = Yes	1	0
Did you receive at any time during the year:/Public Health insurance		[0] = No [1] = Yes	0	0
Did you receive at any time during the year:/Other, specify:		[0] = No [1] = Yes	0	1
Specify your assistance		Various	0	Food voucher
Which food assistance did you receive during the last 6 months?		One or combination of: Food box Freegan, Food box K1, Hot Meal Captains Rib Shack, Hot Meal COME center, Evoucher K1, Evoucher NLRC	Evoucher NLRC	Evoucher NLRC
Which food assistance did you receive during the last 6 months?/Food box Freegan		[0] = No [1] = Yes	0	0
Which food assistance did you receive during the last 6 months?/Food box K1		[0] = No [1] = Yes	0	0
Which food assistance did you receive during the last 6 months?/Hot Meal Captains Rib Shack		[0] = No [1] = Yes	0	0
Which food assistance did you receive during the last 6 months?/Hot Meal COME center		[0] = No [1] = Yes	0	0
Which food assistance did you receive during the last 6 months?/Evoucher NLRC	Most relevant here since this is the evoucher from the NLRC	[0] = No [1] = Yes	1	1
Which food assistance did you receive during the last 6 months?/Evoucher K1		[0] = No [1] = Yes	0	0
During registration we've asked you for your preferred type of food assistance, did you receive food assistance of your preference?		Yes, No	Yes	Yes
What did you prefer?		0, Evoucher, Food parcels	0	0

Do you know on what basis you were selected for the food assistance programme?		Yes, No	Yes	Yes
How do you rate the communication before receiving assistance (time of distribution, location, modality)?		good very good bad very bad	very good	good
What could be improved?	Possibility for writing recommendations		0	0
How do you rate the quality of the distribution (waiting time, information provided during distribution, organization)		0 good very good bad very bad	very good	good
What could be improved?	Possibility for writing recommendations		0	0
Did you feel safe during the distribution?		0 very safe somewhat safe not safe	very safe	very safe
What could be improved?	Possibility for writing recommendations		0	0
Did you feel that the staff was friendly and helpful during the distribution?		0 very friendly/helpful somewhat friendly/helpful not friendly/helpful	very friendly/helpful	very friendly/helpful
What could be improved?	Possibility for writing recommendations		0	0
Did anyone ever asked you to pay anything in order to receive food assistance?		0 No Yes	No	No
Was the money provided on the evoucher enough for your household to buy food and hygiene items?		0 very sufficient somewhat sufficient not sufficient	somewhat sufficient	somewhat sufficient
Did you ever experience any difficulties with using the evouchers in the supermarket?		0 No Yes	No	No
What difficulties did you experience?	Personal clarifications		0	0
What difficulties did you experience?/Card didn’t work		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Lost card		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Forgot pin code		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Forgot 5 letter code		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Took long time to do the		[0] = No	0	0

transaction		[1] = Yes	0	0
What difficulties did you experience?/Return to supermarket because of wrong transaction		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Supermarket Staff was not helpful		[0] = No [1] = Yes	0	0
What difficulties did you experience?/Other (specify)		[0] = No [1] = Yes	0	0
Specify your difficulty:	Personal clarifications	Various	0	0
Did you experience any increase in prices in the supermarkets since you started using the voucher?		0 No Yes	No	Yes
How much did prices increase		0 small increace some increase large increase	0	some increase
What items did you mostly buy with the evoucher?		One or combination of: Meat Beverages Vegetables Canned food Other (specify) Cleaning/hygiene items Dry food (rice, legumes)	Vegetables Beverages Dry food (rice, legumes) Meat Cleaning/hygiene items	Vegetables Dry food (rice, legumes) Meat
What items did you mostly buy with the evoucher?/Vegetables		[0] = No [1] = Yes	1	1
What items did you mostly buy with the evoucher?/Beverages		[0] = No [1] = Yes	1	0
What items did you mostly buy with the evoucher?/Dry food (rice, legumes)		[0] = No [1] = Yes	1	1
What items did you mostly buy with the evoucher?/Canned food		[0] = No [1] = Yes	0	0
What items did you mostly buy with the evoucher?/Meat		[0] = No [1] = Yes	1	1
What items did you mostly buy with the evoucher?/Cleaning/hygiene items		[0] = No [1] = Yes	1	0
What items did you mostly buy with the evoucher?/Other (specify)		[0] = No [1] = Yes	0	0
Specify the items you mostly bought with the evoucher		Various	0	0
Do you have any suggestions on how to improve the evoucher assistance?		Yes No	No	Yes
Please give your suggestions:		Various	0	If it can be extended cause not working
How would you rate the quality of the meals?	Not relevant		0	0
What would you like to change	Not relevant		0	0

Were the portions sufficient for you?	Not relevant		0	0
Were the meals enough to meet the food needs in your household?	Not relevant		0	0
Do you have any suggestions on how to improve the daily hot meals?	Not relevant		0	0
Please give your suggestions:	Not relevant		0	0
How would you rate the quality of the food boxes?	Not relevant		0	0
What would you like to change	Not relevant		0	0
Was there enough food provided to meet the food needs in your household?	Not relevant		0	0
Would you like to change the food items that are provided in the food boxes	Not relevant		0	0
What would you like to change	Not relevant		0	0
Do you have any suggestions on how to improve the foodboxes?	Not relevant		0	0
Please give your suggestions:	Not relevant		0	0
How would you rate the quality of the hygiene items?	Not relevant		0	0
Were there enough hygiene items to meet the needs in your household?	Not relevant		0	0
Would you like to change the hygiene items that are provided?	Not relevant		0	0
What would you like to change?	Not relevant		0	0
Do you know how to reach the Red Cross, K1, Freegan, Captains Rib Shack or COME Center when you have questions or complaints?		Yes No	Yes	Yes
Did you ever contact the Red Cross, K1, Freegan, Captains Rib Shack or COME Center?		Yes No	Yes	No
How did you contact us?		Call by Phone Other (specify) Send a WhatsApp Sent an email Visit the office	Call by Phone Send a WhatsApp	0
How did you contact us?/Call by Phone		[0] = No [1] = Yes	1	0
How did you contact us?/Send a WhatsApp		[0] = No [1] = Yes	1	0
How did you contact us?/Visit the office		[0] = No [1] = Yes	0	0
How did you contact us?/Sent an email		[0] = No [1] = Yes	0	0
How did you contact us?/Other (specify)		[0] = No [1] = Yes	0	0
Specify how you contacted us:		Various	0	0
When you contacted Red Cross, K1, Freegan, Captains Rib Shack or COME Center, was the problem solved (timely and with desired result)		Yes No	Yes	0
Please specify your problem:		Various	0	0
Please click validate to finalise your survey. Thank you for filling out the survey and providing us with valuable feedback.		0	0	0

Do you have any other suggestions on what we can do better?		Various	No everything is ok just would like a little more on the card	Everything is good so far just would really like it to extend more till I Get a job cause things hard out there.
Thank you for your time		0	0	0
_id	Personal ID		18	25
_uuid	User ID		e23e5c8c-a5b8-40a7-b7b2-0f001eb03933	9685f411-e7a2-4e61-8e19-ff02086c52ce
_submission_time	Time		2020-12-08T16:07:11	2020-12-09T13:13:52
_validation_status	Empty	0	0	0
_notes	Empty		[]	[]
_status	Submitted	submitted_via_web	submitted_via_web	submitted_via_web
_submitted_by		0 collect_sxm	0	0
_tags	Empty	0	0	0
_index	Number of entry	1 to 73	1	2

D.3. BNFnonPII (Geo-demographics)

Demographics of all the beneficiaries who participated in the eVoucher, Hot Meals and/or Food Boxes project. 'BNF_nonPII' means beneficiary non personal identifiable information. The dimensions of this dataset are the following: 59 columns (around 55 potentially relevant) and 6367 rows (around 5507 potentially relevant).

Table D.3: Description of geo-demographic registration data

Column name	Explanation	Example 1	Example 2
id	Unique ID for every beneficiary (representing household)	5f217241c10162ff0	5f217241c4fba8d01
deleted	Not useful	0	0
description	Possibility for beneficiaries to leave a message.	My family is thankful for the voucher because am not working and it's very hard to provide	NULL
contact_phonewa1	Telephone number with whatsapp available	yes	yes
contact_phone1notown	Telephone is own possession.	yes	yes
address_neighborhood	Neighborhoods (67 unique)	Union Farm	Simpson Bay Village
address_district	Districts (11 unique)	Lower Prince's Quarter	Simpson Bay
finance_rent	Housing situation (homeless, not_own_not_pay, own, own_mortgage, pay_rent, pay_rent, (blank)	pay_rent	pay_rent
finance_salary	Receiving salary (yes, no, (blank))	yes	yes
finance_o_a_v	Receiving elderly pension	no	no
finance_s_s_r_p	Receiving Sint Maarten Stimulus and Relief Plan	no	no
finance_onderstand	Receiving social safety net from the Government (Onderstand)	no	no
finance_salary_amount	Amount of salary	1200	1500
finance_salary_amount_currency	Currency of before mentioned (ANG or USD)	ANG	ANG
finance_a_o_v_amount	Amount of elderly pension	NULL	NULL
finance_a_o_v_amount_currency	Currency of before mentioned (ANG or USD)	ANG	ANG
finance_s_s_r_p_amount	Amount of Sint Maarten Stimulus and Relief Plan receiving	NULL	NULL
finance_s_s_r_p_amount_currency	Currency of before mentioned (ANG or USD)	ANG	ANG
finance_onderstand_amount	Amount of social safety net receiving	NULL	NULL
finance_onderstand_amount_currency	Currency of before mentioned (ANG or USD)	ANG	ANG
finance_total_income	All income and relief combined	1200	1500
finance_total_income_currency	Currency of before mentioned (ANG or USD)	ANG	ANG
modality_preference	Aid modality preferred (eVoucher, Food Box, Hot Meal, (blanks)). Filled in by beneficiary.	eVoucher	eVoucher
modality_assigned	Aid modality assigned (eVoucher, Food Box, Hot Meal, (blanks)). Filled in by aid organization	eVoucher	eVoucher
modality_supermarket	How many people in the household are able to go to supermarket? (NULL, 0, 1, 2, 3, 4, 5, 6, 7, 8, 11, 15)	1	1
modality_cook	How many people are in the household who can cook? (NULL, 0, 1, 2, 3, 4, 5, 6, 7, 8, 15)	1	1
modality_vegetarian	Preference for vegetarian food boxes or hot meals. (yes, no, (blanks))	no	no
composition02	Number of people with the age between 0 and 2 in the household	1	0
composition317	Number of people with the age between 3 and 17 in the household	3	1
composition1864	Number of people with the age between 18 and 64 in the household	3	2
composition65	Number of people from the age 65 and upward in the household	0	1
composition_pregnancy	Pregnant women in the household? (number, yes, no, (blank))	no	no
composition_pregnancy_amount	How many pregnant women are in household? (NULL, 0 or 1)	NULL	NULL
composition_total_h_h members	Total members of household	7	4

composition_under_aged_ho_h_h	Number of persons in household under 18	0	0
medical	Are there medical issues in the household?	yes	yes
medical_other	Are there other than earlier mentioned medical issues in the household? (0, 1)	0	0
medical_other_s_pecify	Possibility to note other medical conditions	NULL	Medication asperin
medical_allergies	Allergies? (0 = No,1 = Yes)	0	1
medical_diabetes	Diabetes? (0 = No,1 = Yes)	0	0
medical_high_cholesterol	High cholesterol? (0 = No,1 = Yes)	0	0
medical_high_blood_pressure	High blood pressure? (0 = No,1 = Yes)	0	0
medical_disabled	Disabled? (yes, no, (blank))	no	no
medical_disabled_amount	Number of disabled persons in household	NULL	NULL
rr_i_d	Unique ID (connected to NLRC Inventory Transaction List)	ACIPB	ABYLJ
status	Is filled in by staff members of aid organisation. Most relevant are: active, noAnswer (invited to participate but did not show up), cancelled (was part of previous phase but not anymore), non-eligable (was not selected to participate because of criteria), Eligible follow-up programme, PH3: non-eligible (only for phase 3), investigation (needed extra investigation, mostly because of duplicates in one household), Planned for distribution (invited for distribution but did not show up), Disconnected phone (no contact details), No data check (no data available)	No data check	active
eligibleph1	Eligible for phase 1? No, because of other aid or pension. (Yes, no_other, No_highincome, no_pension, no_onderstand, no_ssrp, No_french)	Yes	Yes
eligibleph2	Eligible for phase 2? (yes, no_income, no_onderstand, no_french)	yes	yes
implementingpartner_id	(5f10f3bc9d1df3f12 = Red Cross, 5f19a0318afeb87dc = K1 Direct, NULL, 5f16ee907af9c716c = Freegan, 5f19a09fefd40a55f = Captains Rib Shack, 5f19a0b598db62de4 = COME Center)	5f10f3bc9d1df3f12	5f10f3bc9d1df3f12
language	Language spoken (English, Spanish, Creole, French)	English	Spanish
statuscancelled	If canceled, explanation why by staff members.	NULL	NULL
prim_ben_d_o_b	Date of birth of primary beneficiary	NULL	NULL
prim_ben_gender	Gender of primary beneficiary of household (male, female, not_say)	male	female
householdsize	Household size (small, medium, large)	large	medium
partof_phase1	Part of phase 1	1	1
partof_phase2	Part of phase 2	1	1
topupamountmontly	Monthly amount of aid on eVoucher (167 USD, 250 USD, 333 USD, NULL)	333 USD	250 USD
contactchannelexport	(WhatsApp, SMS)	WhatsApp	WhatsApp
phase3	Part of phase 3 (0 = No,1 = Yes)	0	1

D.4. NLRC Inventory Transaction List

The Netherlands Red Cross Inventory Transaction list exists of all the transactions of vouchers. Every transaction is one between a voucher and supermarket or between a voucher and program. There are 13 columns (around 10 relevant) and 184.406 rows (92.203 potentially relevant, this is halve of the dataset since every transaction exist of two rows).

Table D.4: Description of voucher data

Column name	Explanation	Example 1	Example 2
Id	ID number of row. Goes from 1 to 184.406	1	2
Date	Date and time of transaction	May 25 2020, 00:04:41	May 25 2020, 00:04:41
Commodity	Type of commodity. Only E-USD in this dataset.	E-USD	E-USD
received	Money received on the voucher of beneficiary	0	128
spent	Money spent at the supermarket	128	0
Amount	Transaction between voucher and supermarket	-128	128
Unit/Currency	Currency (in this dataset there is no deviation from USD)	USD	USD
membertype	Could be Program, Beneficiary or Shopkeeper	Program	Beneficiary
member_id	ID related to program or shopkeeper	c32ccff5-b883-42f4-acc0-9c51285b41d2	0
Name	Name of transactions. When membertype is Program this will be named Supermarket Vouchers. When membertype is Shopkeeper this will be named the specific supermarket.	Supermarket Vouchers	0
Id	ID	-	AADTA
Tag	Tag related to beneficiary	-	5f4911a9c3d101eed
Category	Category can be Activity, Beneficiary or Vendor.	Activity	Beneficiary

D.5. Receipts

During the voucher project on Sint Maarten, there was an attempt in collecting all the receipts from supermarket visits. An estimated 10 percent of all receipts has been collected. Most of these receipts are generated by the supermarkets Carrefour Bush Road and Fairway. Around 534 receipts from the Carrefour Bush Road from a time frame of one month (1st October until 27th of October) is available. The receipt data can be read as text via software (so no conversion is needed from picture to text). The amount on the receipts is in Netherlands Antillean guilder (NAF) instead of US Dollar. This research did not use these receipts for market basket analysis and customer segmentation, but this could be applied in further research. Figure D.1 gives an example of a receipt from Carrefour Bush Road.

```
Charge A/C (NAF)
31818
AUTHORIZATION:
 A M O U N T :   2 4 . 0 5
SIGNATURE ______________________________
CASHIER NAME: ARIANNA
C0085    #7633    14:25:30    27OCT2020
            S00079    R004
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Charge A/C (NAF)
31818
AUTHORIZATION:
 A M O U N T :   2 4 . 0 5
PLEASE KEEP FOR YOUR RECORDS
CASHIER NAME: ARIANNA
C0085    #7633    14:25:30    27OCT2020
            S00079    R004
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
BABY FOOD
  CRF BB LINGETTES 2X7             6.04
   PROMOS                         -1.09
NET PROMOTION PRICE ->>            4.95
  CRF BB LINGETTES 2X7             6.04
   PROMOS                         -1.09
NET PROMOTION PRICE ->>            4.95
  CRF BB LINGETTES 2X7             6.04
   PROMOS                         -1.09
NET PROMOTION PRICE ->>            4.95
GROCERIES
  MC FS CLOVES WHOLE               5.79
NON-FOOD
  *PAPER BAG CHARGE                0.36
PRODUCE
  *APPLE RED
 0.565 kg @      5.40/ kg          3.05
  6 BALANCE DUE                   24.05
    Charge A/C (NAF)              24.05
      [K] 31818
    RED CROSS 2020
    CHARGE ACCOUNT
    1721-55318
    CHANGE                         0.00
YOUR SAVINGS TODAY!
-------------------
  TOTAL DISCOUNTS        3         3.27
----------------------------------------
      SIGN UP FOR MEMBER CARD
Using a member card on this transaction
would have given you 20 Points. Please
check with our customer service for
details.
       Our Card Program gives you
         upto 10% Voucher back.
----------------------------------------
CASHIER NAME: ARIANNA
C0085    #7633    14:25:30    27OCT2020
            S00079    R004
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Figure D.1: Receipt (example from Carrefour Bush Road)