



**University of
Zurich^{UZH}**

DEPARTMENT OF INFORMATICS

Informatics and Sustainability Research Group

MASTER THESIS

Design, Development and Evaluation of Interactive Dashboards for the SafeConsume Project

Author:

Tim Grimm,
16-706-285,
tim.grimm@uzh.ch

Supervisors:

Prof. Dr. Lorenz Hilty
hilty@ifi.uzh.ch

Prof. Dr. Joachim Scholderer
joachim.scholderer@uzh.ch

JUNE 13, 2022

Abstract

Research projects often generate large amounts of data that are of interest not only to the scientists who were originally involved in the project, but also to other scientists and various groups of stakeholders. However, most open-data initiatives only enable access to raw data, requiring considerable data analysis expertise and implicitly limiting the prospective users of the data to scientists with such expertise. In order to make open-data accessible also to stakeholder groups outside the scientific community, other forms of data representation may be required. The master project documented in this thesis uses dashboard techniques for this purpose and a research project funded by the EU under the Horizon 2020 program as a case. In Study 1, the development of four interactive dashboard prototypes is documented from requirements gathering to launch. In Study 2, the usability of the dashboards is assessed in terms of task completeness, task completion time, the System Usability Scale (SUS) plus a qualitative analysis of think-aloud protocols collected from test users. Taken together, the results show that the created dashboards are easy to use and that the visualized research data is understandable for users with and without prior knowledge about the research project that served as a case. Possible extensions to dynamic data and semi-structured verbal data are discussed.

Keywords: Dashboard Development Process, Data Visualization, Usability Testing, Research Data, SafeConsume, Interactivity, Design Guidelines, Tableau Software

Acknowledgements

I would like to thank Prof. Dr. Lorenz Hilty for giving me the opportunity to write my master's thesis at the Department of Informatics at the Informatics and Sustainability Research Group at the University of Zurich (UZH) and for being my advisor.

Moreover, I would like to thank Prof. Dr. Joachim Scholderer for being my supervisor during the whole master's thesis. I dearly acknowledge your guidance, assistance and valuable advice during my whole project.

A special thanks also go to all the interview partners who participated in my experiment. Without them, this work could not have been done.

Furthermore, I would also like to thank Gajendra Sivajothi, who helped me with all the technical issues at the beginning of the project. Finally, yet importantly, I would like to dedicate thanks to my beloved family and close friends for their unconditional support and inputs throughout the whole period of this project.

Contents

1	Introduction	1
1.1	Background	1
1.2	Objectives	2
1.3	Overview	3
2	Theory and Previous Research	4
2.1	Dashboards as a Data Visualization Approach	4
2.1.1	Definition	4
2.1.2	Purpose	4
2.1.3	Development Process	5
2.1.4	Design Guidelines	6
2.2	Usability of Dashboards	9
2.2.1	Definition	9
2.2.2	Evaluation Methods	10
2.2.3	Application of Evaluation Methods in Practice	15
2.2.4	Results of Studies	16
3	Research Questions and Case Context	19
4	Study 1: Dashboard Development	23
4.1	Development Process	23
4.2	Requirements Engineering	27
4.3	Results	30
4.3.1	Relative Risk Dashboard	32
4.3.2	Contribution to Risk Dashboard	34
4.3.3	Impact Potential Dashboard	34
4.3.4	Risk Communication Dashboard	36
4.4	Deployment	38
5	Study 2: Usability Evaluation	39
5.1	Method	39
5.1.1	Participants	39
5.1.2	Procedure	39
5.1.3	Analysis	41
5.2	Results	43
5.2.1	Task Completion	44
5.2.2	User Responses	46
5.2.3	SUS Questionnaire	51
6	Discussion	54
7	Conclusions and Recommendations	61
8	Bibliography	63
9	Appendix	69

List of Tables

1	Detail Statistics about the Top Five Foodborne Hazards in Europe (SafeConsume, 2022, Hazards)	20
2	Comparison of Achieved Task Completion Degree in Iteration 1 and 2	46
3	Results of the Content Analysis of the Think-Aloud Protocols: frequencies of topics by dashboard independent of iteration stage of the dashboards	47
4	Singular Value Decomposition(SVD) Results	47
5	Mean System Usability Scale Scores for Dashboards after Testing Iteration 2	53
6	Usability Test Participants List	70

List of Figures

1	Structure and Research Areas of the SafeConsume Project (SafeConsume, 2022, The Project)	21
2	Suggested Retail-to-Consumption Process (SafeConsume, 2022, The Project)	22
3	Development Process Steps for Creating and Implementing a Dashboard from Requirement Gathering till the Launch of the Dashboard	23
4	Design Thinking Process used for Ideation Step (Zarnegar, 2019)	25
5	Consistent Design of Dashboard Header including Logo, Title, Filters and Info Icon	30
6	Example of Filter Function	31
7	Example of Info Icon and the Additional and More Detailed Information that appear	31
8	Example of the Hover Over Feature in a Visualization	31
9	Relative Risk Dashboard	33
10	Contribution to Risk Dashboard	35
11	Impact Potential Dashboard	36
12	Risk Communication Dashboard	37
13	Mean Task Completion Times (in Seconds) as a Function of Dashboard and Testing Phase (Error Bars indicate 95% Confidence Intervals)	45
14	Correspondence Analysis Biplot of Content Analysis Results independent of Iteration Stage of the Dashboard	48
15	Mean SUS Score as a Function of Dashboard and Testing Phase (Error Bars indicate 95% Confidence Intervals)	52
16	SUS Questionnaire for Dashboard Testing	72
17	Test Environment Website: Starting Page for Usability Testing	73

Acronyms

BI Business Intelligence

DALY Disability Adjusted Life Years

ISO International Organization for Standardization

NGO Non-Governmental Organizations

UI User Interface

UX User Experience

RQ Research Question

SME Small and Medium Enterprises

SUS Software Usability Scale

SVD Singular Value Decomposition

UZH University of Zurich

WP Work Package

1 Introduction

1.1 Background

Nowadays, data is collected everywhere, and in the last year 79 zettabytes of data were generated worldwide (Djuraskovic, 2022). A zettabyte is a unit of measurement for storage capacity and is equal to a thousand exabytes, a billion terabytes or a trillion gigabytes (Indicative, n.d.). This is already a massive amount of data and forecasts say that the amount of generated data constantly increasing year by year. By 2025, more than 150 zettabytes of Big Data will need to be analyzed (Djuraskovic, 2022). Not only the amount of data is increasing, also the investments in research projects within Europe dealing with such data is raising. A current example is the ninth framework program "Horizon Europe" by the European Framework Programmers for Research and Innovation. It runs from 2021 to 2027 and has a total budget of around EUR 95.5 billion, which is EUR 10 billion more than the last framework program called "Horizon2020" (SERI, 2021). And wherever research is conducted, data is often collected as well. In order to complete a research project, the collected data and the gained knowledge must somehow be communicated. One option would be to simply store the raw data in a database and make it accessible to future interested parties. However, since the data is often very complex and therefore difficult to understand for outsiders, it takes a lot of time and knowledge to interpret the data and derive information from it. Therefore, a more efficient way to make research data more understandable and accessible to everyone should be considered.

A way to do this is using visualizations methods which help people understand and interpret data more quickly and thoroughly (Srinivasan, Drucker, Endert, & Stasko, 2018). Dashboards are particularly well suited for visualizing information, as they aim to show only the most important data and insights at once (Few 2006). In addition, they offer the possibility to develop them interactively to give future users more flexibility in exploring the data. When developing a dashboard, it must be taken care to ensure that the dashboard is intuitive while maintaining a high level of

user friendliness (Calzon, 2021). This allows users to take advantage of all the dashboard's features and presented information. The goal is to give them a good overview of the research data and its context. Therefore, it is advantageous to test the dashboard with possible end users before publishing it (Romanelli, 2020). This way, even complex research data can be made understandable and accessible to everyone.

1.2 Objectives

The goal of this thesis is to design, develop and evaluate interactive dashboards for the data of a research project. It is expected that the whole process follows state of the art principles and uses the latest technologies. The following objectives build the scope of the project:

Design and development of dashboards:

The purpose of this thesis is to design and develop some interactive dashboards which present and summarize the gained knowledge through the research project in the most efficient and understandable way for future research in the same area. The end users should quickly get used to the dashboards without a comprehensive application guide.

Usability evaluation of dashboards:

In addition, the dashboard should be evaluated using a range of different methodologies with the goal of presenting actionable insights on their design, functionality, and usability. This will allow future researchers and other interested parties to quickly get used to the dashboards. Furthermore, it makes the results understandable and interpretable which facilitates further research in this area.

Provide simple maintenance and further development possibilities:

In the end, interactive dashboards should be provided that are easy to maintain and accessible for everyone via the internet. It is also important that interested users have the possibility to further develop the dashboards on their own and that updated data or newly obtained data can easily be

incorporated.

1.3 Overview

The structure of this thesis is as follows: First, the Chapter 2 introduces the concepts of using dashboards as a data visualization approach and the most important guidelines that needed to be considered when designing a dashboard. It then examines the usability of dashboards and analyzes studies conducted on the usability of dashboards and present the obtained results. Then in Chapter 3 the research questions and the underlying research case context, the SafeConsume project, is presented.

The main part of this thesis is divided into two studies: Study 1 in Chapter 4 shows the whole process of the dashboard development phase from the requirements gathering until the launch of the dashboards. The created dashboards are presented and explained. Study 2 in Chapter 5 is about the usability evaluation of the dashboards. The used evaluation method is explained and the evaluation results are presented.

In Chapter 6, the main findings that emerged from the two studies are discussed. Whereas, Chapter 7 concludes the outcome of this thesis and also gives some recommendations for future work in dashboard design and dashboard usability testing.

2 Theory and Previous Research

This chapter introduces the concept of using dashboards as a data visualization approach and what should be looked for when designing a dashboard. It then examines the usability of dashboards and analyzes studies conducted on the usability of dashboards and presents the obtained results.

2.1 Dashboards as a Data Visualization Approach

The aim of this section is to create a clear and thorough understanding of the dashboard concept and its characteristics following a discussion of the development process of dashboards and some design guidelines.

2.1.1 Definition

Many different types of definitions for dashboards exist in literature (Wexler, Shaffer, & Cotgreave, 2017). Few (2006) defines dashboards as visual displays of the most important information which are needed to achieve a user's objectives, all arranged on a single screen so that the information can be monitored at a glance. According to Yigitbasioglu and Velcuc (2012) a dashboard is a graphical user interface (UI) which provides measures of business performance for managerial decision making. Pauwels et al. (2009) describe a dashboard as a relatively small set of interconnected key performance metrics and underlying performance drivers that reflect both short- and long-term interests to be viewed across the organization. Although these provided definitions varies, some commonalities can be found, and the dashboard concept can be understood as a compact visual presentation of important information.

2.1.2 Purpose

First let's start with a well-known and often seen example of a dashboard's purpose in literature: The dashboard in a car. It consists of different visual indicators to inform the driver about the status of the car's various systems. The dashboard's purpose is to ensure the correct functioning of

the car and it helps the driver to achieve his goal, which is driving from point A to point B. When something is wrong on the way, the driver is notified about the current problem and is thus able to take action (Janes, Sillitti, & Succi, 2013). If this example is translated into a business setting, it follows that dashboards in organizations are used to monitor their performance (Yigitbasioglu & Velcu, 2012). Various sources of data are gathered to provide a complete picture of the organization's performance. Meaning that managers do not have to comb through multiple reports to find pieces of information and put it together themselves to get the whole picture of the situation. By providing relevant, timely insights into how the business is going, rapid and adequate responses and measures to problems are possible (Dover, 2004). This facilitates the decision-making process in an organization (Eckerson, 2010). Furthermore, in some cases, dashboards can be used to perform data analyses (Yigitbasioglu & Velcu, 2012) by having the ability to quickly access valuable information without needing the assistance from the IT department to retrieve this information from a database (Dover, 2004). In general, dashboards have the goal to inform the user without distracting them from their tasks (Few, 2006). To reach that goal, dashboards use charts, tables, gauges, and other kind of visualizations methods in order to show the information in the simplest possible and most comprehensible way (Janes et al., 2013).

2.1.3 Developement Process

The dashboard development process is like any other interface design process, whereas the only main difference is that the outcome is a dashboard (David, 2021). Commonly the process begins with requirements gathering by asking the right questions to the future users and stakeholders to get a deep understanding of the accurate needs of the dashboard. It is about understanding the type of requests, finding out the target audience and what they want to see on the dashboard as well as which data sources should be used and how often the dashboard should be refreshed. This is in all design processes the most important step (Chakrabarti et al., 2011). It forms a solid basis for the further development and will strongly influence the next steps (Pillai, 2020). In the next step, first dashboard mock-ups are created, which is mostly done by sketching out ideas with a pen on paper (David, 2021). It is about determining which charts to use, where to locate them

on the dashboard and how the whole single screen will be structured including title, logos, icons, and metrics. At this stage, feedback can already be obtained from users, stakeholders, or the team members. Besides all the initial designing, it is also important to plan through the whole project's tasks as well as the following steps. Thereafter, the needed data is extracted from different data sources, transformed, and loaded into the backend of the future dashboard. During this step it is important to reach a high data quality which is an indispensable requirement. Moreover, as many calculations and transformations should already be done in order to save time and complexity in a later stage as they are not needed to be performed in the Business Intelligence (BI) tool anymore (Pillai, 2020). Once the prototype is satisfying and fulfills all requirements as well as the data is in the desired structure, the dashboard is created (David, 2021). When the dashboard is built, it should be tested for functionality and usability before it is published. The functional testing ensures that all included features, like filters or navigations, work properly and that the data is displayed correctly (Pillai, 2020). After checking that the dashboard is working as planned, user friendliness is also tested. This testing step is to ensure that the dashboard is intuitive enough so that end users understand it within a few seconds and do not need any additional documentation like a user manual (Romanelli, 2020). However, for internal purposes, user- and technology-related documents, such as information about the data sources, a factsheet about the dashboard or a user manual, should be created (Pillai, 2020). Once the dashboard has been fully tested and documented, it can finally be launched to the desired parties (David, 2021).

2.1.4 Design Guidelines

As already mentioned earlier, dashboards visualize data and ideally doing this in a useful way in order that the users can fulfill their goals. Unfortunately, many dashboards are not designed in the right way. They are thus not considered consumer friendly and therefore fail to communicate efficiently and effectively. To prevent this a dashboard must be designed by following at least a few design guidelines and practices that are aligned with the way people see and think (Few, 2006). Although there is no definitive number of guidelines, many of them cover the same critical areas. In the following section the most important guidelines are summarized.

Target Audience: A dashboard is usually designed for someone else than the one designing it. Meaning that the developer needs always to keep in mind who will view and use the dashboard and customize it to that target person or target group (Few, 2006). Because the users are the ones who decides whether the system will be used and considered as easy to interact with or not (Schooley, Hilton, Abed, Lee, & Horan, 2011). This guideline is generally important for designs, not just for the design of dashboards.

Screen Boundaries: A user should be able to get the full overview of all the displayed information at one glance. Thus, it is important that a dashboard fits within one screen and no content it is hidden (Few, 2006). Calzon (2021) also explicitly mentioned that smaller screen sizes, like the one from smartphones, should be taken into consideration when a dashboard is designed. The user should have the same user experience (UX) in terms of the available information and possible functionalities when using the dashboard on smaller screens compared to when using it on a bigger screen, like a desktop.

Display Essential Data: Since the size of the dashboard is limited, a dashboard should only display the most essential data needed for analysis and decision making. This is different to a report where a lot of extra information is included (Few, 2006).

Data Needs Context: Data on a dashboard must always come with some background information. In the absence of context, the user cannot tell if the data he is considering is good or bad, and whether any action needs to be taken or not (Calzon, 2021).

Data Layout: When the dashboard is designed, human visual perceptions should be considered. Thus, it should be decided which data should be grouped together and where the data should be displayed. It is recommended that the most important data should be placed at the top left of the dashboard (Calzon, 2021; Few, 2006).

Choosing the Display Media: The presented data needs to be displayed using the correct display media, whether it is by a text, a graph, or a map. Do not experiment with different display medias just to provide variety, use the media that conveys the data most effectively (Calzon, 2021; Few, 2006). It can be hard to find the correct type of chart for the data and maybe the user has a different opinion on it, so it is important to consider it carefully and test it in order to get feedback.

Highlight the Important Information: Few (2006) divides the data on the dashboard into two categories. The first category is data that it is always important and the second category is data that is only important in the moment. For the second type of data, highlighting is an effective way to attract the users' attention right away. However, highlighting data should only be used sparingly, as overusing it might defeat its purpose (Asif, 2016).

Choosing the Color Palette: Using too many colors can confuse the user. So, the number of colors should be limited and used in a consistent way. Using several shaded of one color is recommended however, it is very important that they can easily be distinguished (Calzon, 2021).

Make the Dashboard Attractive: This section includes a lot of the above-mentioned guidelines: The dashboard and all its elements should look aesthetically pleasing. This includes the color choice, high resolution on text and images, content layout, font type and other style elements (Few, 2006). According to Calzon (2021) the modern dashboard should look minimalistic and clean. In other words, flat design is trendy nowadays.

The design guidelines mentioned above help the developer to design a dashboard. Summarized a few principles are important in order to create, design and implement good dashboards: The developer should always have in mind to keep the dashboard as simple as possible and really understand the data behind it. In addition, it is important to know the used visualization tool as well as the target audience. If all these principles are taken into account by the developer, it is already a good basis for developing a usable and easy understandable dashboard.

2.2 Usability of Dashboards

Dashboards should be easy and intuitive for the user. Hence, usability testing and evaluation is an essential part of the dashboard development process (Few, 2006). In the following subsections, the term usability is defined, methods for usability evaluation of dashboards are presented and usability testing examples from practice are included.

2.2.1 Definition

Usability represents a quality feature of a product and stands for the usability of the product (Jordan, 2020). The International Organization for Standardization (ISO) has defined usability under the standard ISO 9241 for ergonomics of human-system interaction as «*extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*» (ISO, 2018). The three keywords, namely efficiency, effectiveness, and satisfaction occur in every usability definition (Dumas & Redish, 1999; ISO, 2018; Jordan, 2020). Effectiveness describes the «*accuracy and completeness with which users achieve a certain goal*». Efficiency is defined as «*the effort used to achieve a certain goal in relation to accuracy and completeness*», whereas satisfaction is defined as «*freedom from annoyance and positive attitudes toward the use of the product*» (ISO, 2018). For Dumas and Redish (1999) usability means that users can use a product quickly and easily and thus achieve their goal, which can be compared to Few's (2006) purpose of a dashboard. A developer of a dashboard should always have in mind that the focus is on the user's side, which needs the dashboard to be productive. The user is always busy and just wants to accomplish its tasks as effective and efficient as possible with the highest degree of satisfaction as possible. Thus, the user is the one who decides whether a product is easy to use or not. In addition to the standard ISO definition, learnability of a system can be seen as a very important aspect of usability as well. It is about how easy it is for new users to learn the dashboard's functionality and perform a certain task (Rubin & Chisnell, 2008).

2.2.2 Evaluation Methods

There are various methods for evaluating usability. In the following sections only the most used methods and techniques for evaluating dashboards in studies are listed. This contains heuristic ratings (Nielsen, 1994), usability testing (Dumas & Redish, 1999), focus groups (Kuniavsky, 2003) and questionnaires (Brooke, 2013; Dumas & Redish, 1999).

Heuristic Evaluation

In the heuristic evaluation, a set of usability principles, called heuristics, is tested for compliance. These heuristics point to certain problem categories in the design of interactive systems and describe desired characteristics of the interaction between the user and the system (Nielsen & Molich, 1990). This procedure is performed by one or more usability experts who try to put themselves in the shoes of the user. The main goal of a heuristic evaluation is to uncover potential usability problems (Lavery, Cockton, & Aktinson, 1997). Among the very commonly used and most recognized principles regarding this evaluation method are Nielsen's (2020) ten heuristics, which are designed to help avoid the most important common problems in user interface design. These 10 heuristics for evaluating an interactive system are listed and described below:

1. Visibility of system Status: *«The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time»* (Nielsen, 2020).
2. Match between system and the real world: *«The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order»* (Nielsen, 2020).
3. User control and freedom: *«Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process»* (Nielsen, 2020).

4. Consistency and standards : *«Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions»* (Nielsen, 2020).
5. Error prevention: *«Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action»* (Nielsen, 2020).
6. Recognition rather than recall: *«Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.»* (Nielsen, 2020).
7. Flexibility and efficiency of use: *«Shortcuts - hidden from novice users - may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions»* (Nielsen, 2020).
8. Aesthetic and minimalist design: *«Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility»* (Nielsen, 2020).
9. Help users recognize, diagnose, and recover from errors: *«Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution»* (Nielsen, 2020).
10. Help and documentation: *«It is best if the system does not need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks»* (Nielsen, 2020).

Usability Tests

Usability tests are another well-known method in dashboard usability testing. The crucial difference to the heuristic evaluation is that in this method the product under test is used. In a usability test, people are observed solving various tasks with the help of the dashboard. Usability tests are also particularly suitable for evaluating the usability criteria of effectiveness and efficiency. Since the participants use the dashboard directly during the tests, it is very easy to observe how efficiently and effectively people can perform tasks with it (Dumas & Redish, 1999). Usability tests have the following five characteristics:

- The goal is to improve the usability of a product: The main goal of this type of testing is to improve the usability of the tested product. This already distinguishes it from other types of tests, such as functional tests, where the main objective is to ensure that the product conforms to the specification. Often the tests also aim to check the usability criteria of efficiency, effectiveness, and satisfaction (Dumas & Redish, 1999).
- Participants represent real users: Another important characteristic is that the participants of the test come from the actual target group for the product. This means that it should include those people who already use or will use the product (Dumas & Redish, 1999).
- Participants perform real tasks: Just as important as the use of real users is that these people carry out real tasks in the context of the tests that they would also have to carry out in real use of the product. This is the only way to uncover problems that really get in the way of the work. To do this, however, it is necessary to understand the task area and the field of application of the product to subsequently be able to derive suitable tasks. In the case of very extensive or complex systems, this can be difficult for outsiders and in some cases, it is only possible to test some part of the real tasks (Dumas & Redish, 1999).
- One observes and records actions and statements of the participants: The basic idea of a usability test is to observe the participants during the execution. Depending on the type of test, this can happen in different ways. To enable further analyses and evaluations, conspic-

uous behaviour, verbal statements and comments as well as other data must be recorded. In addition to carrying out tasks, participants can be explicitly asked for their opinion, for example in the form of a questionnaire, in order to collect further data (Dumas & Redish, 1999).

- One analyzes the data, determines the real problems, and recommends changes to solve these problems: After a test, the collected data is evaluated together with the user's own observations in order to derive concrete usability problems. Based on this, suggestions for improvement can be made to remedy these problems (Dumas & Redish, 1999).

To ensure that the right thing is tested and that the test runs as smoothly as possible, it is necessary to plan the test and its components. Otherwise, important aspects of the product and its use may not be tested, problems may occur during the test, or the data collected may be incorrect or incomplete. Therefore, the objectives, participants, tasks, test environment and material as well as the data to be collected must be considered and precisely defined in advance (Albert & Tullis, 2013).

As already described in the characteristics of a usability test, it is crucial that the dashboard is tested with real users. In the planning stage it is therefore necessary to identify the current and potential users and thus recruiting appropriate participants according to the defined target group. In the selection process, it is important to consider those characteristics of the people that are shared by all members of the target group, but also those that may represent differences and define subgroups within the target group. For example, the target group could include both novice and experienced users. The more characteristics are considered, the more subgroups can be identified, the more detailed the test. For a test, however, it is usually not possible and often not relevant to include people from all possible subgroups. Therefore, the subgroups with the highest relevance in relation to the goals and which are expected to reveal the most important usability problems should be selected. For example, in a first usability test one would include the typical users who make up the largest proportion, whereas in a repeated test one might be more interested in a specific subgroup. Basically, it can be assumed that the more people participate, the more different

problems occur, but the most serious and important problems can be observed again with each person (Dumas & Redish, 1999). According to Nielsen and Landauer (1993), about 80% of usability problems can be detected with only five users.

Once the participants have been selected, the tasks can be defined based on the goals for the use of the product and the participants' own goals for the usability test. It is of great importance that the scenarios described within the tasks correspond to the real world of the participants (Barnum, 2010). Even before a test is carried out, it must be defined which data should be collected or need to be recorded during the test and which measurements are necessary for this. If you then want to make statements about the individual usability criteria, you must also consider in advance what concrete data can be used to measure efficiency, effectiveness, and satisfaction (Albert & Tullis, 2013).

Focus Group

Focus groups are another way to evaluate a dashboard. This involves bringing together a group of potential users, with a moderator leading the discussion and the participants discussing the product and giving feedback (Kuniavsky, 2003). As it is about understanding the users' desires, motivations and values, this is often done in the early stages of the dashboard development process. The gained inputs help to define the requirements of the dashboard and support the further development of the dashboard (Dowding et al., 2018).

Questionnaires

With the help of questionnaires information about the user's opinion of the usability of a system can be obtained. Instead, of asking the question during the testing. Questionnaires have the advantage to be consistent, which enables the comparability of the answers. Furthermore, the scope is limited using rating scales as an answering option (Dumas & Redish, 1999). The survey of the participants can take place at various times during a usability test. Depending on the information you want to obtain, there are three common times: Before the test, after a task is completed or

after the test (Barnum, 2010).

There exist some standard questionnaires that can be used to determine the satisfaction of users and their perception of the usability of a product. Even if these questionnaires are not tailored in detail to the product under test, the use of standard questionnaires has the advantages of repeated use between various products and the results are still comparable (Sauro & Lewis, 2016). The Software Usability Scale (SUS) questionnaire by Brooke (1996) is one of the most widely used and widespread usability questionnaires for dashboards (Nitsch, Adamcik, Kuso, Zeiler, & Waldherr, 2019; Schall et al., 2017; Wu et al., 2019). Its purpose is to measure the users' subjective perception of the usability of a dashboard. The SUS questionnaire consists of ten sentences, half of which are positive and half negative. For each sentence, the user can express his or her level of agreement or disagreement on a scale of five points (Brooke, 2013).

2.2.3 Application of Evaluation Methods in Practice

In order to test the usability of a dashboard, the presented evaluation methods are often combined in practice. This can be seen in multiple studies which are analyzed in the following section (Hamm, Money, Atwal, & Ghinea, 2019; Martinez et. al, 2018; Nitsch et al., 2019; Schall et al., 2017; van Lieshout et al., 2016; Wu et al., 2019).

Dowding & Merrill (2018) evaluated the usability of a clinical dashboard using heuristic evaluation and task analysis. The heuristic evaluations were conducted by experts. The experts received a checklist which consisted of the ten general usability principles by Nielsen (2020). To every principle some questions like "Does every screen have a title or header that describes its content?" were asked and the experts investigated whether this was the case or not. The subsequent question was then answered accordingly with yes or no. In addition, the experts were also given some tasks that covered the full functionality of the dashboard. The time was measured and later compared to normal user's time spent for completing the same tasks. The smaller the difference between the two groups, the more the dashboard was seen as usable.

Compared to Dowding & Merrill (2019) most of the studies focus on user feedback rather expert feedback and therefore do not use any heuristic evaluation (Danny et al., 2019; Dowding et al., 2019). In the study of Danny et al. (2019), a mixed-methods design was used to evaluate the usability of an interactive dashboard. It included a series of semi-structured interviews and a usability testing procedure using a think-aloud protocol as well as a SUS questionnaire. Also Lagha et al. (2020) make use of task-based evaluation, semi-structured interviews and a validated usability satisfaction questionnaire for testing the dashboard. For the task-based evaluation standardized tasks are used, covering as much of the dashboard's functional scope as possible. During the task performance time and completeness is measured. It is also striking that the participants are always asked to think aloud while completing the tasks and maneuvered through the dashboard to gain some qualitative data as well. The SUS survey is applied as the most standard and validated tool for testing the usability of dashboard through a questionnaire. This has been the most observed approach in several usability studies for dashboards (Hamm, Money, Atwal, & Ghinea, 2019; Martinez et. al, 2018; Nitsch et al., 2019; Schall et al., 2017; van Lieshout et al., 2016; Wu et al., 2019).

2.2.4 Results of Studies

Results of different dashboard usability studies have shown that dashboard elements which belong together should be connected and therefore be grouped together and delineated in a simple way from other elements. Furthermore, it has been shown that consistency not only in the visual appearance of the visualization, but also in the choice of type of visualization has a positive influence on usability. Moreover, it was found that color coding has a significant impact on usability. Using too many light and dark colors can quickly become visually tiring, worsening usability from an end user perspective (Schall et al., 2017; van Lieshout et al., 2016; Wu et al., 2019). Besides the influence of dashboard's design on usability, it has also been seen that users have a different ability to understand data. How good they understand the information on dashboards is often determined by their individual's level of numeracy, which is the ability to understand numerical information, and graph literacy, which stands for the ability to understand information presented

graphically (Gaissmaier et al., 2012). Studies on the visualization of complex health data highlight the importance of considering numeracy and graph literacy of the end users when evaluating data presentation formats. It is shown, that in general, people with low numerical and graphical skills will have significant comprehension problems, regardless of how the data is presented. In contrast, people with high graphic literacy, regardless of their numerical skills, can understand quantitatively presented data like for example in tables, better than people with low graphic literacy (Dowding et al, 2019; Febretti et al., 2014; Garcia-Retamero & Cokley, 2013). Garcia-Retamero and Cokley (2013) have found that it is easier for end users with higher graphics skills and low numerical skills to understand and interpret data that is visually represented in formats such as icon arrays, bar charts or line graphs. Since people have different abilities, dashboard developers should consider providing flexibility in regards of displaying the same information in different ways. It helps to evaluate end users' ability to understand information before a dashboard is published to the target audience (Dowding et al., 2019).

It is known that the inability to understand statistical information is not only because of the end users' ability but also due to the inadequate presentation of the information. Understanding complex research data and metrics, such as relative risk, can lead to confusion and misinterpretation as it is difficult to understand what the percentage refers to. Gigerenzer and Edwards (2003) developed some simple tools for helping users to understand these kinds of metrics. Statistical information can be represented in conditional probabilities (e.g. *«The probability that a woman has breast cancer is 0.8%»*) or natural frequencies (e.g. *«Eight out of every 1000 women have breast cancer»*). Natural frequencies are how people have been exposed to statistical information for most of their history and such information is still easier for most people to process (Galesic, Gigerenzer, & Straubinger, 2009). The same principle can be applied to the understanding of relative risk. Show in an example case: If four out of every 1000 women (age 40 or older) who do not undergo mammography screening die of breast cancer, compared with three out of every 1000 who are screened, then this benefit is often expressed as a relative risk: *«Mammography reduces breast cancer mortality by 25%»*. According to Gigerenzer and Edwards (2003) this relative risk statement can presented

in a more understandable way by doing the following:

- ⇒ Use absolute risks, alone or together with relative risks: *«In every 1000 women who undergo screening one will be saved from dying of breast cancer.»*
- ⇒ Use the number needed to treat or harm: *«To prevent one death from breast cancer, 1000 women need to undergo screening for 10 years.»*

So, instead of using relative risks which results in a confusion in the end user's mind, it has more benefits to use absolute risks alone or in combination with the relative risk or the number of treatments or investigations required to save a life, as seen in the example above. Simple risk representations using natural frequencies clarify the information on a dashboard and can help the end user move from ignorance to insight and facilitate interpretation (Gigerenzer & Edwards, 2003).

3 Research Questions and Case Context

Every research project involves data that is somehow collected over the duration of the project. Most of the data is only available from databases in the form of tables, and it is difficult to interpret it quickly and correctly. In addition, research data is often complex and difficult for an outsider to understand. Thus, this project is looking for a way to pass this data in an efficient and more understandable way to future researcher or interested parties.

A way to do this is using visualizations methods which help people understand and interpret data more quickly and thoroughly (Srinivasan et al., 2018). Dashboards are particularly well suited for visualizing information, as they aim to show only the most important data and insights at once (Few, 2006). In addition, they offer the possibility to develop them interactively to give future users more flexibility in exploring the data. Therefore, overall, the two research questions (RQ) of this project are the following:

RQ I: How should dashboards be designed to visualize complex research data and how should they be made accessible to external stakeholders?

RQ II: How can high usability be achieved and how can it be evaluated?

This research questions are addressed in the following project and set in the context of a current EU-led research project called SafeConsume, in which the Informatics and Sustainability Research Group of the Department of Informatics of the UZH is responsible for the risks assessments and the data management. The SafeConsume project has collected a lot of data over several years through surveys within different European countries, as for example in Denmark, Germany and Spain. The main objective of this project is to reduce the health burden of foodborne diseases. The aim of SafeConsume is to initiate a proprietary and more comprehensive approach to future food safety research, innovation, education, and policy, and to expand the space of opportunities to improve food safety (SafeConsume, 2022, The Project). As this project ends after five years, it is time to put the collected data and lessons learned into a frame that future researchers and interested parties

can quickly understand and address them and moreover further use and process them for their purpose. Since the project covers many different areas, several interactive dashboards will be designed, developed, and evaluated.

Case Context

The EU's research and innovation program named "Horizon2020" funded the SafeConsume project in May 2017. The timeline of the project is set to end by October 2022 (Cordis, 2022). Since the beginning the project is managed by the Norwegian food research institute called Nofima in Oslo (Nofima, n.d.). Nofima coordinates all the 32 project partners from 14 European countries. The partners are mainly research institutes, governments, or non-governmental organizations (NGO) but also include three multinational companies and four small and medium enterprises (SME) (SafeConsume, 2022, The Project).

Table 1: Detail Statistics about the Top Five Foodborne Hazards in Europe (SafeConsume, 2022, Hazards)

Foodborne Hazard	Most Common Sources	Health burden in DALY (per 100'000 People each Year) in Europe	Cases per Year in Europe	Deaths per Year in Europe
Campylobacter jejuni	Poultry and milk	9	> 4'000'000	> 400
Toxoplasma gondii	Vegetables, fruit and meat	6	1'000'000	Unknown
Salmonella enterica	Egg, egg products and poultry	8	> 1'000'000	>2000
Norovirus	Shellfish, bivalve molluscs and crustaceans	4	> 14'000'000	>400
Listeria monocytogenes	Ready-to-eat food	3	> 1000	> 400

The background to the project is the identification of hazards in food that cause about 23 million cases of illnesses and 5000 deaths in Europe each year. It is known that about 40% of foodborne illness outbreaks occur in private households or kitchens. The main reasons are poor hygiene and inadequate heating and cooling at the consumer stage. SafeConsume focuses on the top five food-

borne hazards in Europe, which are responsible for about 70% of the health burden of foodborne disease. In Table 1 above, these top five foodborne hazards in Europe are listed. Also included in the table are detailed statistical information about disease burden, number of cases as well as number of deaths per year in Europe. Health burden is measured in disability adjusted life years (DALY), which stands for the total health burden where one DALY can be thought of as one year of “healthy” life lost (SafeConsume, 2022, Hazards).

From the above stated facts in Table 1, the overall goal of the project is to reduce the health burden of foodborne illnesses. This is done through the development of tools and products, information strategies, education, and food safety policy. These tasks are divided into nine work packages (WP), which are assigned to the different partners of the project. Figure 1 illustrates the distribution of the SafeConsume’s research among the nine different WPs (SafeConsume, 2022, The Project).

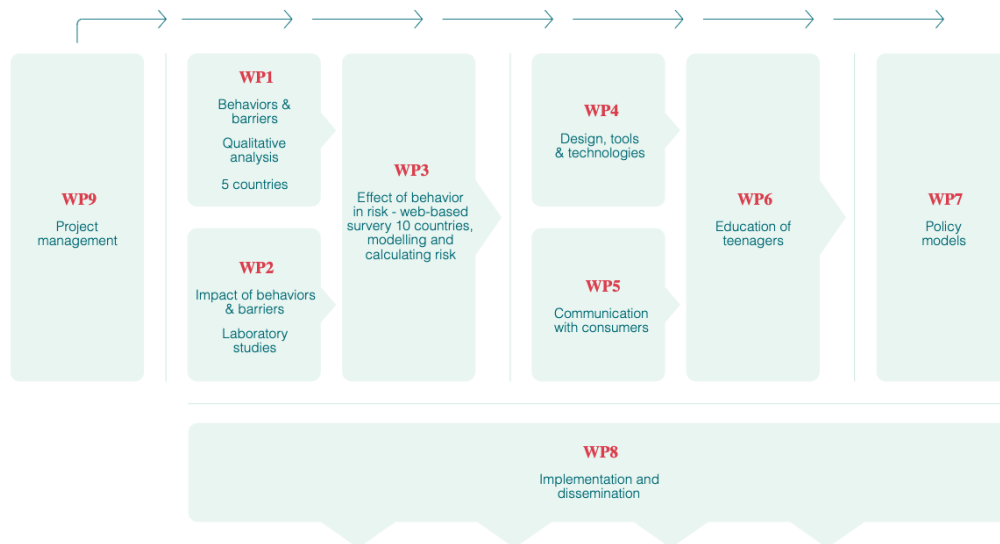


Figure 1: Structure and Research Areas of the SafeConsume Project (SafeConsume, 2022, The Project)

As consumer behavior is defined as both, the core problem and the solution within the SafeConsume project, current food-handling behaviors in all regions of Europe are analyzed through large web-based studies and in-depth investigations. The risk of foodborne infection is examined at

all stages of the consumer process, reaching from retail to consumption, including the selection, storage and preparation of food. The whole process is illustrated in Figure 2. Furthermore, the figure shows what consumers can do differently in order to reduce the chances of illnesses (SafeConsume, 2022, The Project).

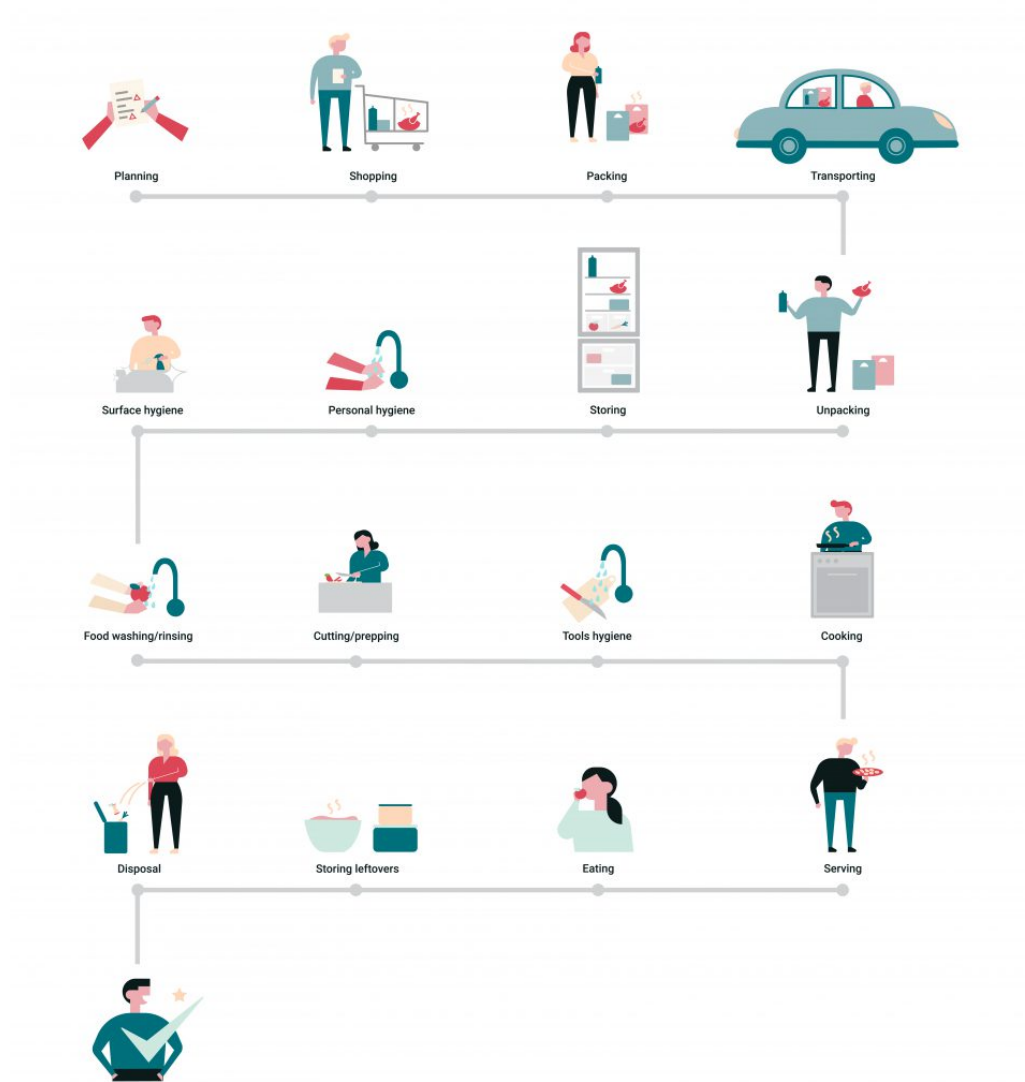


Figure 2: Suggested Retail-to-Consumption Process (SafeConsume, 2022, The Project)

With the insights gained during this project, the SafeConsume project aims to initiate a new and more comprehensive approach in future research, innovation, education, and food safety policy, expanding the space of opportunities for improving food safety (SafeConsume, 2022, The Project).

4 Study 1: Dashboard Development

In this section the method for developing interactive dashboards is discussed and the defined functional and non-functional requirements for the dashboards are listed.

4.1 Development Process

This section characterizes the general development process of the creation and implementation of the dashboards generated in the scope of this thesis. The thesis followed the 7-step process model from Pillai (2020) to build a dashboard, which is pictured in Figure 3. Every step of the model is explained in more detail in the according subsections bellow.

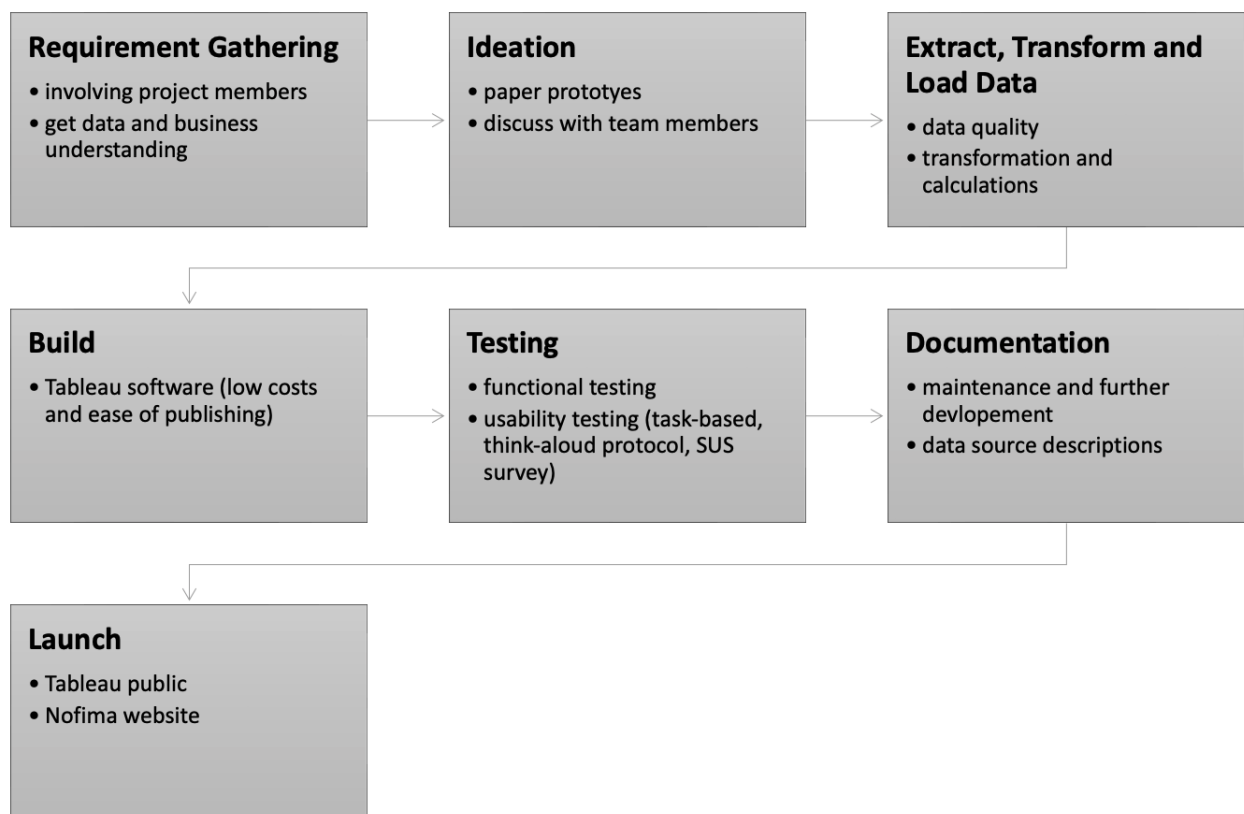


Figure 3: Development Process Steps for Creating and Implementing a Dashboard from Requirement Gathering till the Launch of the Dashboard

Requirement Gathering

In this step, unconventional talks with project members helped to get a deep understanding of the accurate needs of the dashboards. The requirement gathering process started with asking a person from each of the nine WPs. The main question was about the data sources that should be visualized in the dashboards. The SafeConsume has 13 different data set groups. The project members decided that only seven of these data set groups should be made publicly accessible via interactive dashboards. Thus, only the data from these seven data sets served as the fundament for this thesis. Furthermore, reports and raw data from the SafeConsume project were analyzed to build up business and data understanding in order to define further requirements about visualization methods and communication goals. More detailed information about the defined requirements can be found in Section 4.2.

Ideation

After the requirements were defined and a business and data understanding were developed, dashboard-mockups were created. This was mostly done by sketching out ideas with a pen on paper, based on literature and related work about dashboard designing. The sketches set a framework for the charts that should be used as well as where they should be located and how the whole dashboard screen should look like. These sketches were then discussed with other project team members in order to have a different point of views and thus improve the sketches. In this step, the design-thinking approach, shown in Figure 4, was used. This approach allowed many iterations of the "ideation-prototyping-testing" phase to get feedback, incorporate it and therefore further develop the paper prototypes until they were ready for implementation. The "empathize" and "define" steps of the design thinking process were already made in the requirements gathering step and served as basis for the "ideation-prototyping-testing" phase.

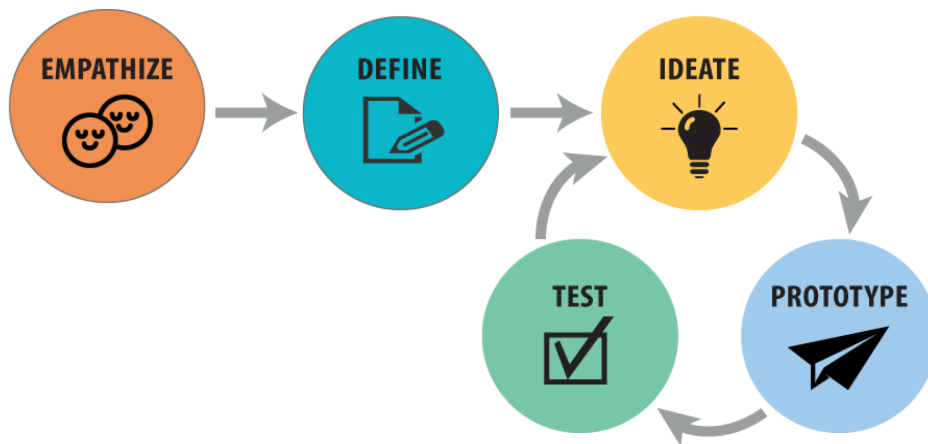


Figure 4: Design Thinking Process used for Ideation Step (Zarnegar, 2019)

Extract, Transform and Load Data

The different data set groups were extracted from different excel files. Subsequently, the data was cleaned and transformed in order to reach a high data quality as well as the right and uniform frame for the future dashboards. The whole data cleaning and transformation step was made with R Studio version 1.3. It is important to mention that the aim of this step was to convert the data into a certain desired form so that there is no need to change the data in any way during the following steps of the creation and implementation of the dashboards. This saves time, storage space and complexity, as the conversion and calculations no longer need to be done in the BI tool.

Build

Once the prototypes are satisfying and fulfilling all requirements as well as the data is in the desired structure, the implementation of the dashboards started. The software for the development of the dashboards was selected based on the requirements of accessibility and maintenance of the dashboards. Main drivers for the selection were maintenance costs, updates possibilities as well as the potential for further development of the dashboards. Regarding these requirements, different software for creating the dashboards, namely the Tableau software, the Power Bi, the Qlik Sense and the Google Data Studio software, were considered, analyzed and compared. The anal-

ysis showed that it makes most sense to use the Tableau software instead of the others. The two main reason for this decision were the upcoming costs of maintaining the dashboards, as well as the ease of publishing the dashboards with this software.

Testing

All dashboards were tested for functionality and usability before being published. The functional testing ensures that all included features like filters or hover functions work properly, and that the data is displayed correctly on every visualization screen. This functional testing was done by testing out every dashboard functionality with project team members. After checking that the dashboard is working as planned, their usability was also investigated. The usability testing was done with eight users including SafeConsume project members and other selected participants from outside the project. The participants were selected as described in Section 5.1.1. The participants were randomly divided into two groups of equal participants, testing the dashboard in two iterations. After the first iteration, the feedback from four participating users was incorporated into the dashboards. The next four participants used the already adapted and thus optimized dashboards for testing. Their feedback was then used for the further and final fine adjustment of the dashboards. More detailed information about the study design of the usability testing can be found in Section 5.1.2.

Documentation

After the completion of this project, the dashboards are handed over to Nofima, the Norwegian food research institute in Oslo. Thus, some additional documentation material was prepared. This contains a documentation for maintenance the dashboard with Tableau and Tableau public, as well as a factsheet for the dashboards with information about the used data sources.

Launch

One requirement of the project was that Nofima has the possibility to integrate the dashboards on their website, which is possible by using Tableau public. This was therefore another strong

reason to use the Tableau software for creating the dashboards during this project. Together with Nofima a Tableau public account was set up and all dashboards were published there internally with Nofima. From there the embedded code for every dashboard can be downloaded and inserted into the according website code, where the dashboards can then be made available for the public. From there, potential users have the option to download the dashboards or just the data behind them. Nofima can also further develop the dashboard by editing the dashboards in the Tableau software and publishing them in an updated version on its public Tableau account. This allows to keep the implementation of iterated or even new and relevant data easy and flexible.

4.2 Requirements Engineering

Initially, requirements engineering was performed based on the seven data sources available for the dashboard implementation. As described above, these data sets were chosen by project members which decided that only these data set groups should be made publicly accessible via interactive dashboards. After several talks with team members, the basic set of features for the dashboards were determined and approved.

The dashboards for the SafeConsume project incorporate the following functional requirements:

- **Performance:** Dashboard components, including the loading of the data, should be as fast and easy as possible in terms of handling.
- **Accessibility:** Facilitating access from all devices with an internet connection was a requirement for the dashboards created within the framework of this project. The dashboards should also have the possibility to be downloaded including their data in order to be used for further development of the data for the sake of own purposes.
- **Responsiveness:** The dashboard should respond to changes in screen size and always aim at using screen estate as effectively as possible.

- **Data:**

- Relative Risk Dashboard: Visualize the risk data from the household survey of WP3.
- Impact Potential Dashboard: Visualize the risk assessment results from the household survey of WP3.
- Contribution to Risk Dashboard: Visualize the purchase-to-consumption process data from the household survey of WP3.
- Risk Communication Dashboard: Visualize the risk communication data from the authority/stakeholder survey of WP7.

- **Visualizations:**

- Relative Risk Dashboard: Implement an interactive dashboard with visualizations of observed risk distribution across EU regions, attributable behavioural risk distribution across EU countries/regions and risks for vulnerable groups.
- Impact Potential Dashboard: Implement an interactive dashboard with visualizations of interventions impact across EU regions, interventions impact on vulnerable groups and information about the intervention itself.
- Contributions to Risk Dashboard: Implement an interactive dashboard with visualizations of the contributions of the total attributable behavioural risk in terms of scope and divided into user journey stages which represent the purchase-to-consumption process.
- Risk Communication Dashboard: Implement an interactive dashboard with visualizations of risk communication across different EU regions and on different user journey stages of the purchase-to-consumption process.

- **User Goals:** The user should be able to work with and explore the data of the dashboards.
 - **Relative Risk Dashboard:** The user can filter the data by the risk assessment endpoints and the region level (country, greater region, region). The user can hover over the different visualizations to get some more detailed information about the regions risk values. The user can sort the values in an ascending and or descending order.
 - **Impact Potential Dashboard:** The user can filter the data by the intervention and the region level (country, greater region, region). The user can hover over the different visualizations to get some more detailed information about the intervention and its impacts. The user can sort the values in an ascending or descending order.
 - **Contributions to Risk Dashboard:** The user can sort the contributions in an ascending or descending order. The user can hover over the different contributions to get some more detailed information about the purchase-to-consumption process and model food.
 - **Risk Communication Dashboard:** The user can filter the data by the journey stage of the purchase-to-consumption process and the material content. The user can hover over the different visualizations to get some more detailed information about the regions risk communication. The user can sort the values in an ascending or descending order.

The following non-functional requirements are met by the implementation of the SafeConsume dashboards:

- **User Interface (UI):** The UI should be clean, usable and visually appealing.
- **User Experience (UX):** The UX should be as pleasant as possible. The UI should give plenty of opportunities for customizations by the user.
- **Maintainability:** The costs for maintenance of the dashboards should be as low as possible. Also, further development of the dashboards should be as convenient as possible.
- **Failure States:** The dashboards should fail gracefully in case of error.

4.3 Results

In this section the developed dashboards for the SafeConsume are presented. First, design elements and features that have been consistently implemented in all dashboards are presented. Then each individual dashboard is shown and set to its context. The links to the dashboards, published on Tableau public, can be found in the Appendix A.

Consistent Design Elements and Features

Each implemented dashboard for the SafeConsume project has the same arrangement of the following components: The logo of Nofima, the title of the dashboard, the filter functions and the info icon are placed on top of the dashboard from the left to the right side in the UI. A line separates these components from all the included visualizations which are placed below this line. This is shown in Figure 5.



Figure 5: Consistent Design of Dashboard Header including Logo, Title, Filters and Info Icon

The same colors were used consistently for text and design elements in all dashboards. Besides the color black, the blue and green tone of Nofima's logo were applied. As already mentioned, if a dashboard includes a filter function, then this function is consistently placed on the top of the UI next to the title. Every filter that can be found within the drop-down menu is a global filter, meaning that this filter applies for all the tables and visualizations shown in the dashboard. The drop-down menu as well as all the filter possibilities are marked in green in order to be immediately noticeable to the user (as shown in Figure 6).

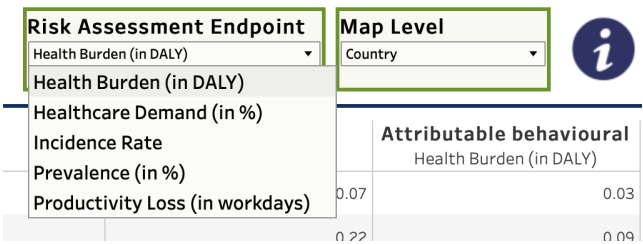


Figure 6: Example of Filter Function

Directly next to the filters there is always a blue info icon. When the user moves the mouse pointer over the icon, additional and more detailed information about the dashboard and its visualizations are displayed (see Figure 7).

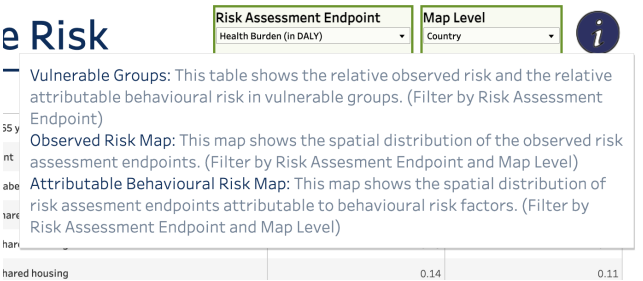


Figure 7: Example of Info Icon and the Additional and More Detailed Information that appear

This hover over function is also applied in every single visualization. It allows the possibility to show the user more information about a certain value. An example of this hover over feature is shown in Figure 8, whereby hovering over a certain region of the map visualization, more detailed information about the value of the selected region is given. As mentioned before, all these described design elements and features can be found across the different dashboards.

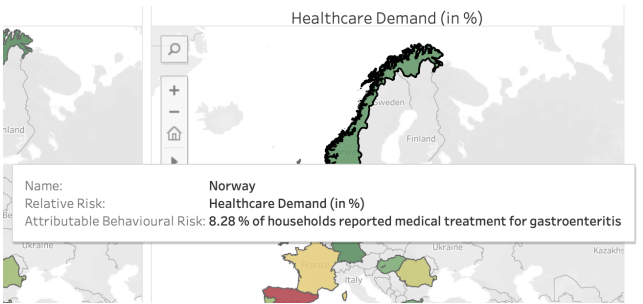


Figure 8: Example of the Hover Over Feature in a Visualization

4.3.1 Relative Risk Dashboard

The Relative Risk Dashboard consists of a table and two geographical visualizations. The data source for this dashboard is based on the household survey of WP3. It is showing the occurrence of an acute gastroenteritis episode in a household in ten countries (Denmark, France, Germany, Greece, Hungary, Norway, Portugal, Romania, Spain and United Kingdom). Different vulnerable groups were selected and defined as followed: at least one household member over 65 years, at least one household member pregnant, at least one household member with diabetes or impaired immune system, men below 30 years living alone or in shared housing, persons 16 - 18 years living alone or in shared housing and persons 19 - 25 years living alone or in shared housing. The information can be filtered in terms of risk assessment endpoints and in terms of the map level, for example on country level. In the following dashboard, pictured in Figure 9, the risk will be assessed in terms of the below listed standard epidemiological risk assessment endpoints:

- Prevalence: Household prevalence of gastroenteritis in % (e.g. 39% of households reported at least one gastroenteritis episode in the past year)
- Incidence: 12-month incidence rate (e.g. 0.55 gastroenteritis episodes per person-year-at-risk)
- Health Burden: Disability-adjusted life years lost (DALY) per 1000 population (e.g. 0.15 DALY per 1000 population)
- Productivity Loss: Workdays lost per household (e.g. 0.62 workdays lost per household)
- Healthcare Demand: Household demand for healthcare services in % (e.g. 9.8% of households reported medical treatment for gastroenteritis)

The dashboard is split into observed risk and attributable behavioural risk. The values for the attributable behavioural risk were estimated by the contribution of people's food consumption and food handling behaviour on top of to the total observed risk. In the table "Vulnerable Group", observed and attributable relative risk is estimated for vulnerable sub-populations. The table can be filtered by risk assessment endpoint and sorted in an ascending as well as a descending order. If

a user hovers over the values more detailed information about its measurement unit is given. The maps in the bottom of the dashboard illustrate the spatial distributions of observed risk and risk attributable to behavioural factors over the regions. The maps can be filtered by risk assessment endpoint as well as by map level. A user has the option of zooming in or zooming out the map as well as clicking or hovering over a region to get more detailed information.

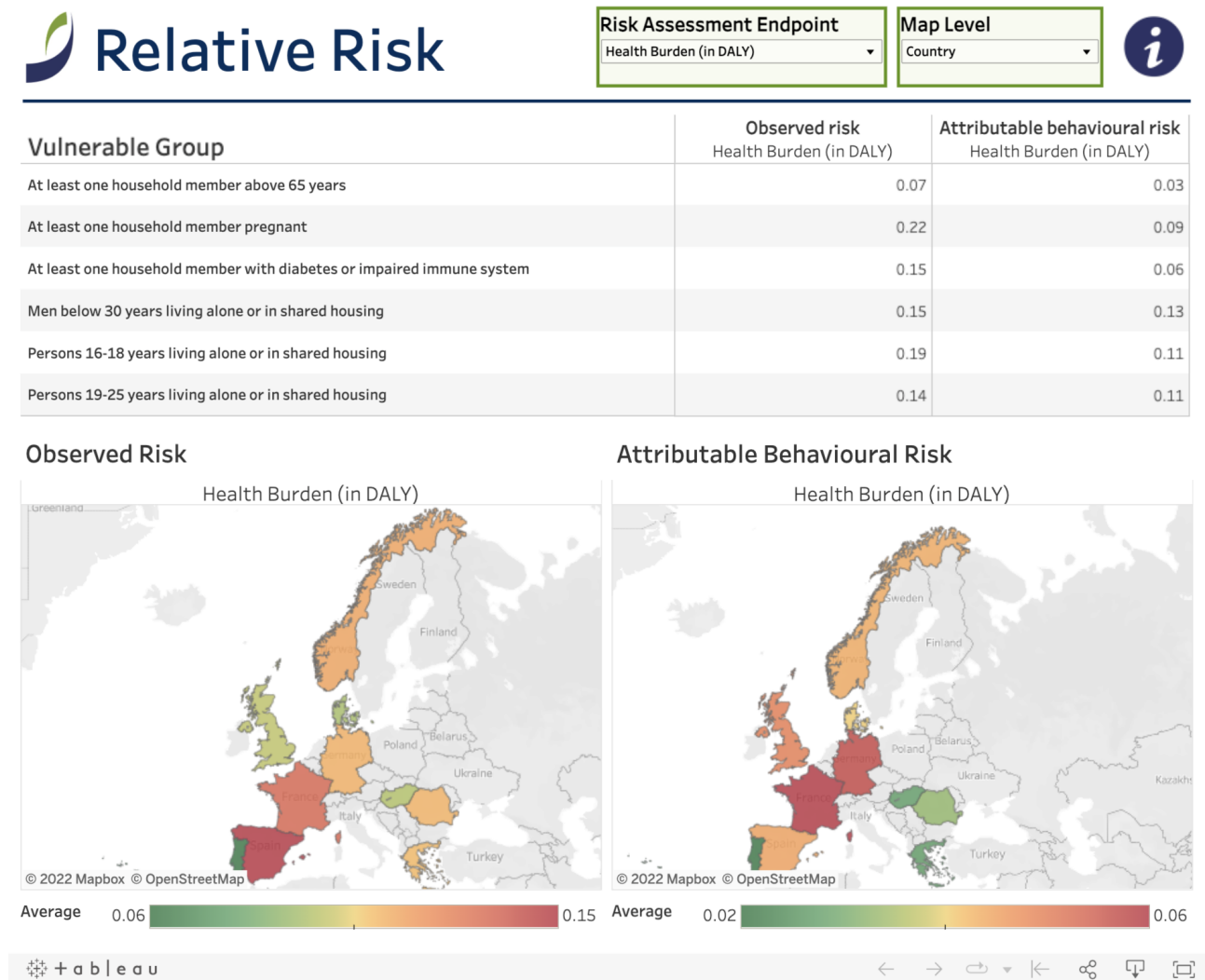


Figure 9: Relative Risk Dashboard

4.3.2 Contribution to Risk Dashboard

The Contribution to Risk Dashboard, shown in Figure 10), consists of one single customized visualization about the purchase-to-consumption processes that is used in the SafeConsume project. A purchase-to-consumption process can be understood as a user journey that consists of different stages in a logical order. The different stages can for example be described as follows: preference and planning, shopping, storing, unpacking, food washing, etc.. The dashboard shows the contributions of the total attributable behavioural risk, cross-classified in terms of scope, meaning the four model food (eggs, chicken, ready-to-eat food and bivalve molluscs) and cross-cutting issues, and divided into user journey stages which represent the purchase-to-consumption process. The stages are defined as: preference and planning, shopping, storing, unpacking, personal hygiene, food washing/rinsing, surface hygiene, cutting/preparing, tools hygiene, cooking, storing leftovers and disposal. So, for example, about 27.8% of the total attributable behavioural risk was associated with cross-cutting issues such as kitchen equipment (including refrigerator use and temperature) and general hygiene. The user has the possibility to sort the visualization in an ascending or descending order, either by stage or only by scope. Moreover, the hovering function is also integrated in this visualization.

4.3.3 Impact Potential Dashboard

The Impact Potential Dashboard, pictured in Figure 11), is about nine possible interventions by which food safety risk in consumer households could possibly be reduced. The nine interventions are: Separate chicken from other food, avoid cutting up raw chicken, sir each surface of a chicken portion, cook food contaminated by chicken, avoid eggs from unsafe sources, store eggs correctly if using raw, check fridge temperature, avoid bivalve molluscs from unsafe sources and eat only well cooked clams and mussels. It is structured into three sections: An info section with a description about the filtered intervention, a table about the intervention's impact on vulnerable groups and a geographical visualization about the intervention's impact in different regions. The info section as well as the vulnerable group table can be filtered by an intervention. The table is sortable in an ascending as well as a descending order. Moreover, if a user hovers over the values,

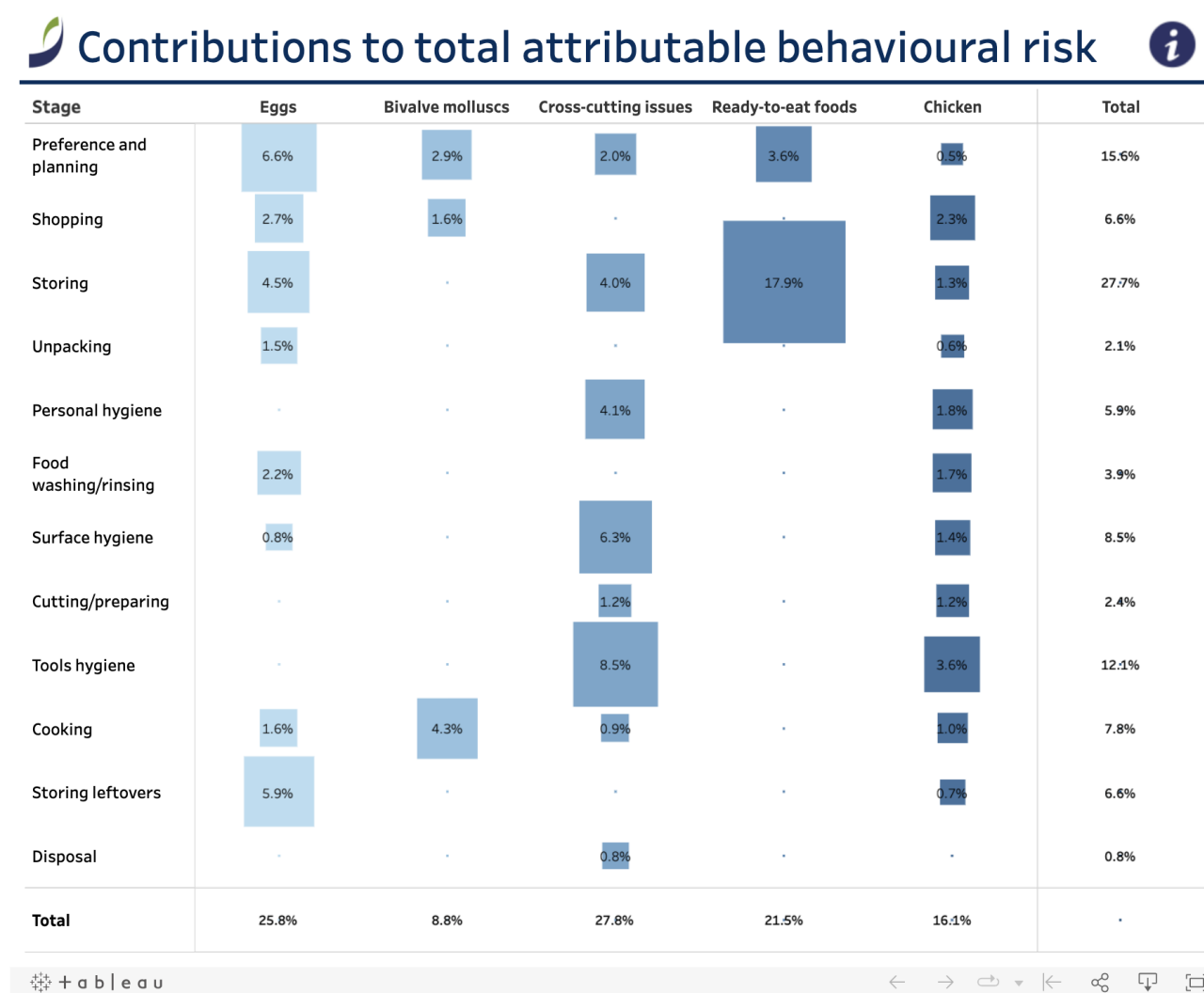


Figure 10: Contribution to Risk Dashboard

more detailed information about the intervention's impact is displayed. The map can be filtered by an intervention as well as by map level. A user has the option of zooming in and zooming out as well as clicking or hovering over a region to get more detailed information. The impact potential of this interventions is defined as the expected reduction in DALY attributable to behavioural risk factors for acute gastroenteritis. The numbers express the risk reduction as a total for the population of the ten countries. Important to mention is that the dashboard below shows the maximum intervention effect that can theoretically be obtained if the intervention reaches 100% of the target population and leads to 100% adoption of the promoted practice in every household of the target population. This important information is shown when a user hovers over the info icon.

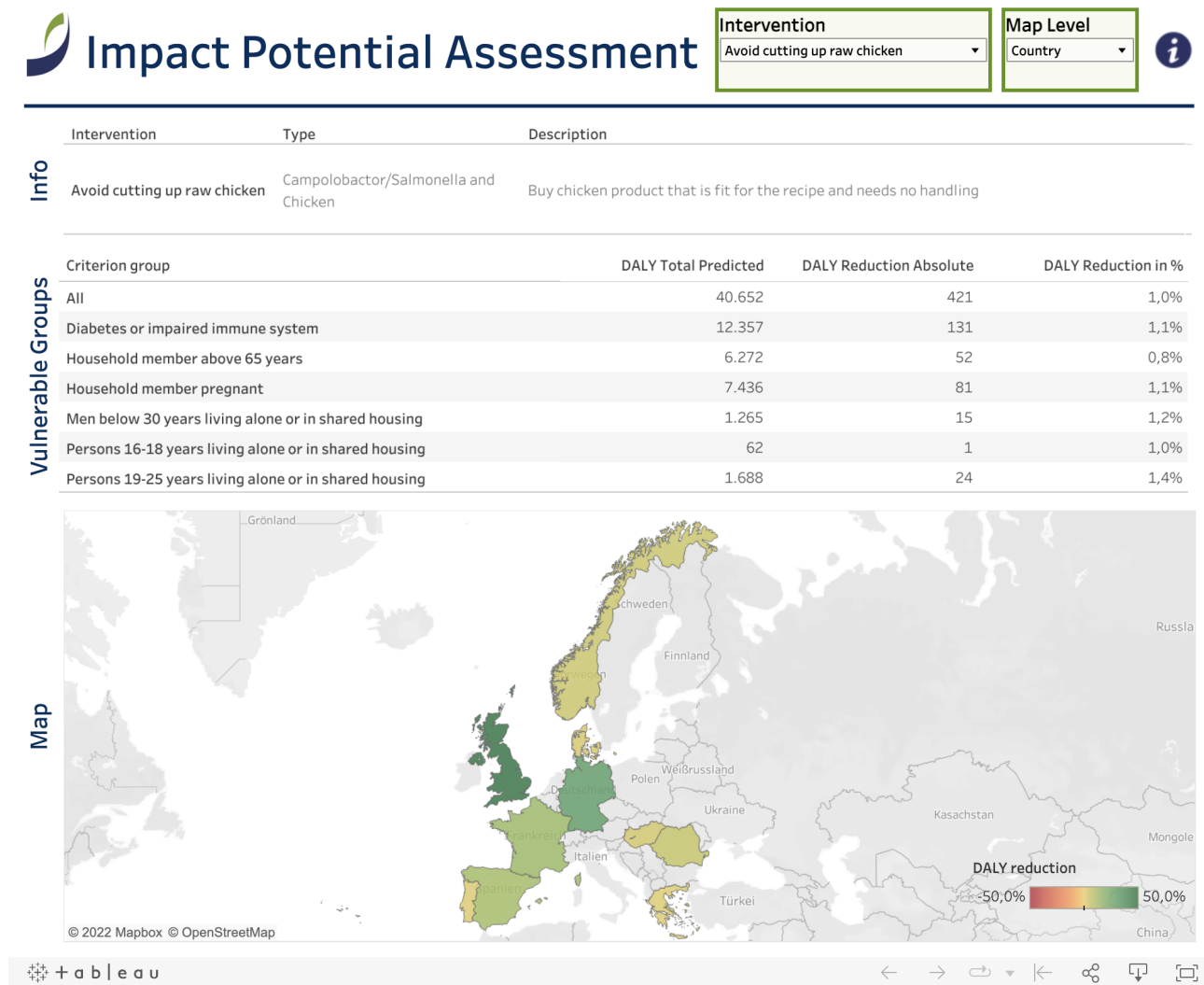


Figure 11: Impact Potential Dashboard

4.3.4 Risk Communication Dashboard

This Risk Communication Dashboard shows the result of a SafeConsume survey about risk communication materials provided by the food safety authorities in different European countries. The dashboard consists of a table and a geographical visualization. To make an impact on the behaviour of the consumer in private households, authorities make use of risk communication tools. These tools help people to make more informed judgements about the food safety hazards and risks they face in their daily lives. The risks were structured based on the stage of the “user journey”, reaching from food choice in the store to reheating of leftovers. Furthermore, depending on the different stages, the focus of the risk documentation was split into different stages, such as

specific foods, most likely pathogens, production or sales locations, etc.. The participating organizations with responsibilities for risk communication towards consumers were asked to indicate possible types of content they provided to consumers in their countries. All of this information is illustrated in the following dashboard, shown in Figure 12. On the left side a sortable table is provided which summarize the risk communication of all countries. On the right side the user can get specific information about a countries risk communication.

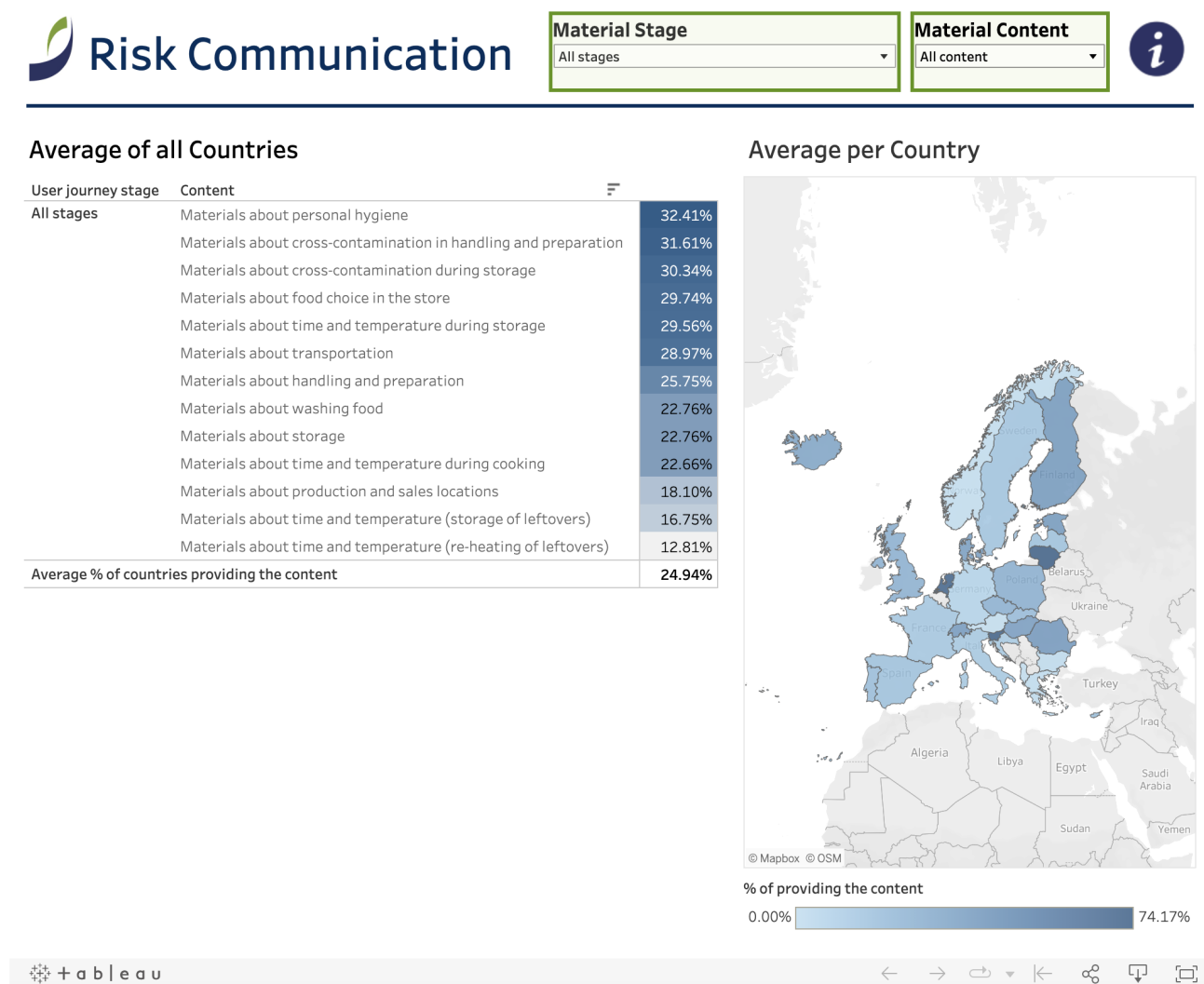


Figure 12: Risk Communication Dashboard

4.4 Deployment

The deployment of the dashboards will be made together with Nofima, since the dashboards will be embedded on their website. Therefore, a Tableau public account is set up for Nofima and all dashboards are uploaded and published there. Tableau public gives the possibility to download the HTML5, CSS and JavaScript code of a dashboard. The code will be inserted into Nofima's content management system and from there embedded into the website code and afterwards published. All used data sets and Tableau worksheet will be stored on Nofima's Tableau public account. From there, potential users have the option to download the dashboards and the data sets behind them. This set up also facilitates further development of the dashboards. Nofima can edit the dashboards in the Tableau software and publishing them in an updated version on their public Tableau account which automatically updates the version of the dashboards on their website, since they are linked to Tableau public. However, going live with the dashboards will be after the submission of this project, latest by the end of august. Nevertheless, meetings have already been held with the ones responsible for the IT-department at Nofima to plan the integration and publishing of the created dashboards on their website.

5 Study 2: Usability Evaluation

5.1 Method

5.1.1 Participants

The participants for the dashboard usability testing were selected from different areas and have never been in contact before with the process of creating the dashboards. However, in general, there can be made a difference between users who are familiar with the SafeConsume project and users who do not know it. From the SafeConsume project three people from different WPs were recruited. They are seen as possible end users of the dashboards. Another four usability testers were chosen because of having a basic knowledge and experience in a project related topic. This means they are familiar with communicating research data or in designing dashboards. Thus, a person with food science, communication and politics, data science and medicine background were recruited. The reason for this heterogeneity within the usability testers was to get broad-based feedback and benefit from the different points of views from people with different backgrounds. Additionally, a user without any knowledge in the mentioned subject areas was also included in the testing to cover the perspective of a totally unknown user. All participants agreed to conduct a 30 to 45 minutes recorded zoom call while sharing their screen. Appendix B gives an overview of all the participating users with information about their familiarity with the SafeConsume project, their inclusion criteria for the testing and the testing iteration they have participated in.

5.1.2 Procedure

The study design included a usability testing method using a think-aloud protocol and a SUS questionnaire. As the focus was on user feedback rather than expert feedback no heuristic evaluation was used in the testing.

Every participating user individually took part in a usability testing session, where they were provided with a link to a website with access to the relevant dashboards. The four, above introduced,

dashboards were tested with each user. All of them tested the same dashboards in the same order. The test was divided into four sections, one for every dashboard. Every section contained an own website including the dashboard, a context belonging to the dashboard, two tasks to solve and the SUS questionnaire to fill out. The procedure for each section was as follows: First, the tester could look at the dashboard and read through the provided context. Then, he was asked to solve some tasks using the dashboard. While doing so, he was asked to think aloud. After completing the two tasks, a structured SUS questionnaire was filled in to collect the participant's feedback. Then the next section started with the same procedure but with a new dashboard. The duration of this individual usability session was around 30 to 45 minutes.

The entire testing was done with eight participants. The participants were divided into two iterations. After the first iteration, the feedback from four participating users was incorporated into the dashboard and thus optimized. Nevertheless, the tasks for the second four users stayed the same. Thus, the next four participants used the adapted dashboards. This was done to investigate whether or not the task performance had improved in the second iteration.

Tasks

For every dashboard two standardized tasks were formulated. Basically, it should only be defined which questions the user should answer based on the data analysis and which information is expected in the result. The task descriptions were carefully worded to mimic realistic information needs from the participating users. Since the focus in this evaluation was on the efficiency and effectiveness in working with the dashboard, the tasks should cover all functionalities and visualizations of the dashboard. The participants were instructed to speak aloud their mind, including what they are thinking, doing, seeing or feeling while they completed a specific task or set of tasks. This helped gathering insights on the difficulties participants encountered while using the dashboards. The list of all tasks can be found in the Appendix C.

SUS Questionnaire

This structured questionnaire was adopted and revised from the SUS, which is a standard and validated tool for usability testing of dashboards (Nitsch et al., 2019; Wu et al., 2019; Schall et al., 2017). As prescribed by the SUS, each participant responded to the same set of 10 questions on a 5-point Likert scale. The used SUS questionnaire can be found in Appendix D.

Test Environment

The participants received a mail with the link to the Zoom call. At the start of the call, the participants were informed about the procedure. Afterwards, they received the link to the testing website in the Zoom chat. The testing website was published via GitHub pages. The link to access the test environment website can be found in Appendix E. All websites were built in a self-explanatory way and the user could easily click through the different dashboard, tasks and questionnaires.

5.1.3 Analysis

The analysis of the usability testing of the dashboards consists of three parts: Qualitative analysis of the user responses, task performance analysis and the analysis of the SUS questionnaires.

Qualitative Analysis

During the tests, think-aloud protocols had been collected from all test users. These interview notes were subjected to thematic analysis to identify emerging themes that may represent strengths or weaknesses of a dashboard. A correspondence analysis was performed to provide a visual and more understandable representation of the results of the thematic analysis. Based on these identified themes and correspondence analysis, some strategies were developed to address weaknesses and improve the dashboard. The results of the qualitative analysis can be found in Section 5.2.2.

Task Performance Analysis

To evaluate the task performance of the users, the completeness of the tasks and the time required for them to fulfill the tasks were calculated. Both the average time to complete a task and the

completeness between the two iterations were compared to determine if the changes made on the dashboards after the first iteration had a positive effect on task performance. A linear mixed model was estimated to determine if there were significant differences between dashboards and test iterations. The results of the task performance analysis can be found in Section 5.2.1.

SUS Questionnaire Analysis

Each test participant completed one SUS questionnaire per dashboard. To analyze the questionnaire, it was necessary to calculate the SUS score, which indicates how good a dashboard is in terms of usability (Bangor, Kortum, & Miller, 2008). For the calculation, the participants' answers were converted into a scale from 0 to 4, with 4 being the most positive answer. The converted responses score for all ten questions were summed up and multiplied by 2.5 as per the scoring instructions, giving a range of possible values from 0 to 100. These calculated mean SUS scores per dashboard were compared within iterations to find possible improvement. Additionally, a linear mixed model was estimated to determine if there were significant differences between dashboards and test iterations. Furthermore, the SUS scores of the second iteration were compared to the benchmark SUS assessment guidelines by Bangor et al. (2008). The results of the SUS questionnaire analysis can be found in Section 5.2.3.

5.2 Results

In response to the testing of the first iteration prototypes, several changes were made in the dashboards. These included:

- Changes to Relative Risk Dashboard:
 - Highlight the filters with green color in tone of the Nofima's logo
 - Adapt the name of the filters in order to make them easier understandable
 - Make fonts of the filter options larger
 - Make fonts of the subsections titles larger
 - Change the color scheme of the geographical map visualizations. Use a more distinguishable color scheme.
 - Add detailed information to the tooltip in the vulnerable groups table
 - Add detailed information to the tooltip for the geographical map visualizations of the observed and attributable risk
 - Add information the about sortable function of the map to the info icon
- Changes to Contribution to Risk Dashboard:
 - Add grey out function if the user clicks on something in the visualization
 - Make fonts of the column and row titles larger
 - Add color levels to the color scheme of the visualization
 - Add information about the sortable function of the visualization to the info icon
- Changes to Impact Potential Dashboard:
 - Highlight the filters with green color in tone of the Nofima's logo
 - Add color levels to the color scheme of the visualization

- Add detailed information to the tooltip in the vulnerable groups table
- Add detailed information to the tooltip for the geographical map visualization
- Changes to Risk Communication Dashboard:
 - Highlight the filters with green color in tone of the Nofima's logo
 - Adapt the name of the filters
 - Make fonts of the subsection titles and filter options larger
 - Adapt the names of the visualization and the subsections
 - Connect the visualizations for mutual filtering by clicking on a data point
 - Add color levels to the color scheme of the visualization
 - Add detailed information to the tooltip for the geographical map visualization

In addition to all these changes, it was discovered that it is important to point out to the user that the dashboards are interactive meaning that the different visualizations can be filtered and sorted. The user should be informed about these possibilities at the beginning of the website's context. The links to the adapted dashboards, published on Tableau public, can be found in the Appendix A.

5.2.1 Task Completion

To assess the overall usefulness of the four dashboards the users had been asked to solve different task using the interactive dashboards. The task completion times were obtained and then statistically analyzed. A linear mixed model was estimated with dashboard (Relative Risk Dashboard, Contribution to Risk Dashboard, Impact Potential Dashboard, Risk Communication Dashboard) and testing phase (iteration 1, iteration 2) as fixed factors and user ID (users 1 to 8) as a random factor controlling for between-user differences in baseline speed. The model was estimated by restricted maximum likelihood and showed a satisfactory fit ($R^2 = .33$, adjusted $R^2 = .28$, $-2 \ln L = 768.18$).

The fixed-effect tests indicated that there were highly significant differences between the four dashboards ($F_{3,83} = 10.09$, $p < .001$). In addition, there were marginally significant effects of the testing phase ($F_{1,5} = 5.35$, $p = .068$) and the interaction of dashboard and testing phase ($F_{3,83} = 2.68$, $p = .054$). In Figure 13, the mean task completion times are plotted as a function of dashboard and testing phase.

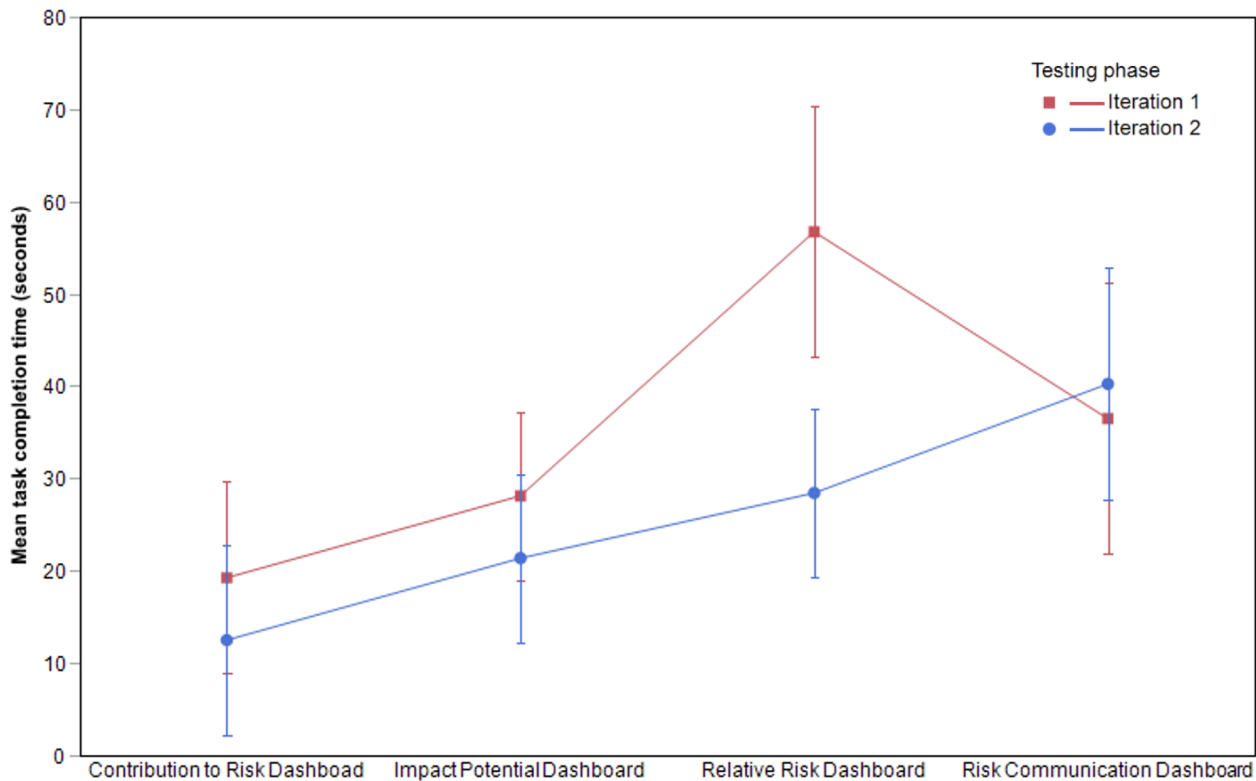


Figure 13: Mean Task Completion Times (in Seconds) as a Function of Dashboard and Testing Phase (Error Bars indicate 95% Confidence Intervals)

The results show that the usefulness of all dashboards, apart from the Risk Communication Dashboard, was improved by the changes made between iterations 1 and 2. The improvement was most pronounced for the Relative Risk Dashboard where most feature and design changes had been made. The Contribution to the Risk Dashboard and the Impact Potential Dashboard also improved in terms of mean task completion time. This is also due to the changes on the dashboards like for example larger fonts in subsections or more color levels on the color scheme of the visualizations. Overall, the changes made it possible for the users to find the answer to the tasks faster and more easily. However, the mean time for task completion in the Risk Communication

Dashboard has deteriorated in the second iteration. But the changes made after iteration 1 positively affected the task completion rate, which can be seen in the Table 2 below. The changes allowed that all users were able to complete every task on the Risk Communication Dashboard which makes the adaptations necessary after all.

Table 2: Comparison of Achieved Task Completion Degree in Iteration 1 and 2

Dashboard	Task Completion Iteration 1	Task Completion Iteration 2
Relative Risk Dashboard	50%	100%
Contribution to Risk Dashboard	100%	100%
Impact Potential Dashboard	100%	100%
Risk Communication Dashboard	75%	100%

Table 2 also shows that the Contribution to Risk Dashboard and the Impact Potential Dashboard already reached a full task completeness in the first iteration, which also explains the smaller changes in the mean task completion time than for the other two dashboards. Finally, it can be said that in the second round all users were able to solve all the tasks, achieving a task completion rate of 100% for all dashboards, which was an important requirement in terms of efficiency of the dashboards and speaks in favour of the changes.

5.2.2 User Responses

During the tests, think-aloud protocols had been collected from all test users. The thematic content analysis yielded altogether 19 topics. A breakdown of their frequencies by dashboard is shown in Table 3.

Table 3: Results of the Content Analysis of the Think-Aloud Protocols: frequencies of topics by dashboard independent of iteration stage of the dashboards

Topic	Relative Risk Dashboard	Contribution to Risk Dashboard	Impact Potential Dashboard	Risk Communication Dashboard
Filters not obvious	3	0	0	0
Filters badly named	1	0	0	3
Interactivity not obvious	1	1	0	0
Table sorting function not obvious	0	2	0	0
Font too small	1	0	0	0
Incomprehensible information	2	0	2	1
Too much hidden information	0	0	0	2
Bad color scheme	4	1	1	0
Difficult to navigate	1	0	3	0
Instructions missing	2	0	0	1
Bad tooltip	1	0	1	2
Bad titles	3	0	0	2
Well structured	0	3	4	0
Easy to use	0	6	2	2
Interesting information	1	3	5	3
User friendly	0	5	3	1
Comprehensible information	0	3	1	0
Good color scheme	0	0	3	0
Nice visualization	0	4	0	0

To obtain a more visual and better-comprehensible representation of the data, the table was subjected to correspondence analysis. The singular value decomposition (SVD) is shown in Table 4. The coordinates of topics and dashboards on the first two dimensions (together accounting for more than three quarters of the total variation) are plotted in Figure 14.

Table 4: Singular Value Decomposition(SVD) Results

Dimension	Singular Value	Inertia	Contribution to Chi ²	Percent	Cumulative %
1	0.75	0.57	49.32	45.48	45.48
2	0.63	0.40	34.83	32.12	77.60
3	0.53	0.28	24.30	22.40	100.00

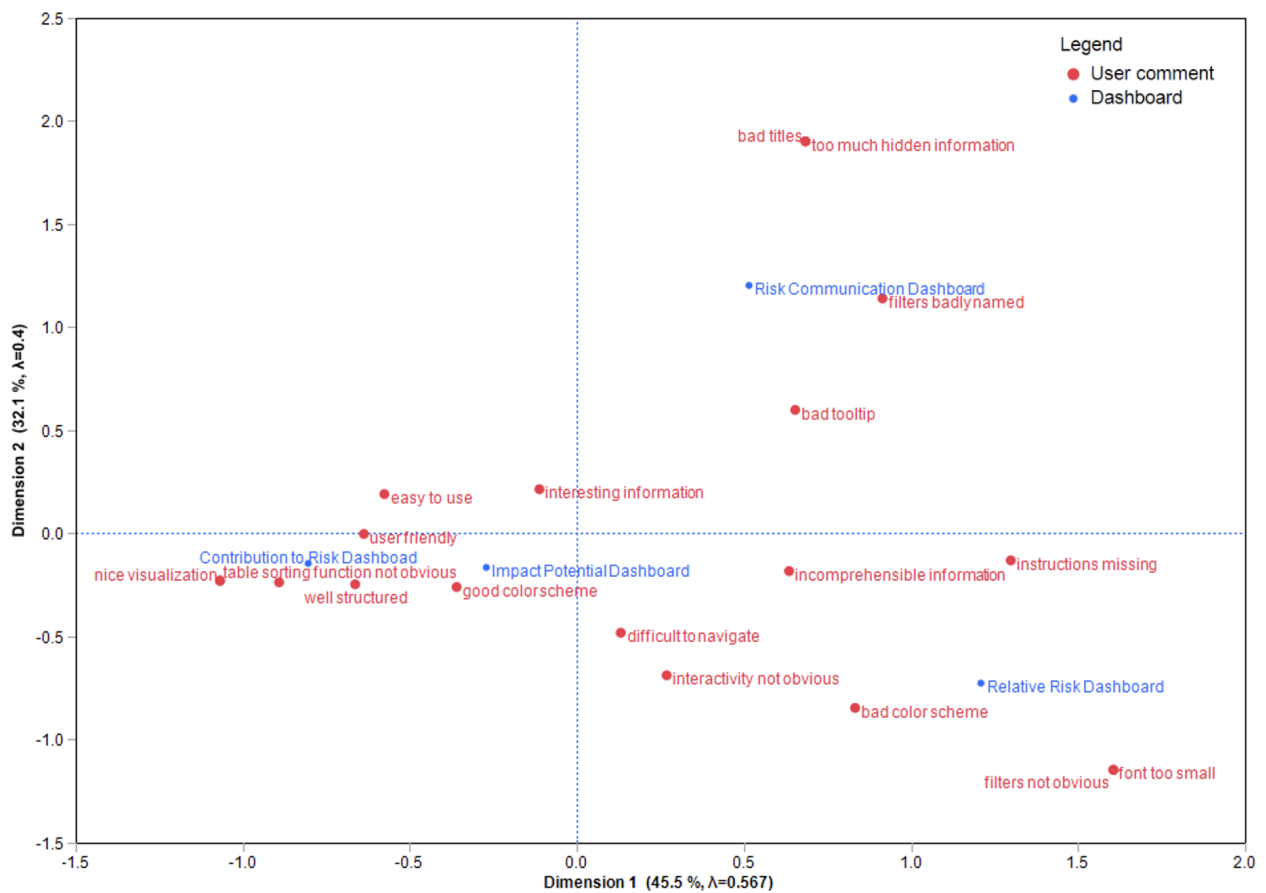


Figure 14: Correspondence Analysis Biplot of Content Analysis Results independent of Iteration Stage of the Dashboard

The first dimension, accounting for 45.5% of the total variation, contrasts positive user responses (such as “nice visualization” and “well structured”) with negative user responses (such as “font too small” and “filters not obvious”). The second dimension, accounting for 32.1% of the total variation, contrasts responses to textual/semantic dashboard elements (e.g. titles) with responses to visual/syntactic elements (e.g. fonts, colors).

The Contribution to Risk Dashboard was mainly associated with responses highlighting the quality and structure of the visualization, but also that the sorting function for the table was not obvious in the dashboard. User 4, for example, commented:

«If I would have known that the visualization is sortable, I could have solved the tasks faster.»

For this reason, the users should be informed in advance that the dashboards are interactive, and

that the visualization can be sorted in an ascending or a descending way. But most of the feedback to this dashboard was positive. User 2 really liked the layout of the visualization:

«The layout of the visualization is nicely designed, because the different size of the squares show in an understandable way at first glance very fast where the important information is.»

Additionally, user 1 had an interesting idea to improve the interaction with the following visualization:

«It would be nice if I click on a field in the visualization, then detailed information would be displayed and all other parts of the visualization would be greyed out.»

This suggestion was integrated after the first testing iteration because it makes the user dashboard interaction more effective. Next to the Contribution to Risk Dashboard, the Impact Potential Dashboard is placed on the correspondence analysis biplot. Here the users liked the good separation of the different dashboard elements. As user 6 for example commented:

«This dashboard looks clean, and the tables and visualization are nicely separated from each other through the font titles on the side.»

Some of the testing participants also mentioned during solving the task that the displayed information is interesting. Furthermore, the color scheme of the geographical map visualization was also praised by different users. User 3 compared it to the color scheme of the geographical map visualization of the Relative Risk Dashboard:

«The color coding of this map was chosen much better than in the last (relative risk) dashboard. Here it is much clearer to distinguish different values by their colors»

User 3 was not the only one who criticized the color coding of the geographical map visualization in the Relative Risk Dashboard. The different shades of blue were considered too indistinct and therefore the colors were adapted after the first iteration. Since the Relative Risk Dashboard was the first dashboard that came up during the usability testing procedure, some of the users mentioned in the first iteration that the interactivity of the dashboard was not obvious. In addition, two users could not solve the tasks of the Relative Risk Dashboard because for them the filter function

was not obvious for them. User 1 made a comment to prevent this in the future:

«For me it was not obvious that the dashboard is interactive and that there is a filter option on the top right corner of the dashboard. Maybe the filters should be highlighted, or the user should be instructed somewhere about the dashboard's interactivity and all integrated features»

Due to this statement, an adjustment was made after the first iteration. The filters on every dashboard were highlighted with a border around it. The color for the border line and background color of the filters were chosen to be in the same green tone as of the Nofima logo. In addition, it is recommended to include a short general instruction for the users about the dashboard's interactivity and features on the Nofima website, before presenting the first dashboard. Furthermore, another weakness of the Relative Risk Dashboard was the difficulty to interpret the displayed information. User 4 mentioned that the values, which were odds ratios, in the vulnerable groups table are hard to understand:

«I never had to deal with odds ratios and so I do not understand exactly what these values indicate.»

User 4 was not the only one who dealt with this problem. Therefore, the values were changed from odds ratios to relative risk values. Furthermore, the tooltip information was extended by adding context to the relative risk values. The bad tooltip was also mentioned in reference to the Risk Communication Dashboard. This has also been adapted there by providing more information to certain relative risk values. Furthermore, user 3 was confused about the naming of the filters:

«The naming of the filter irritates me. I was thinking that 'map content' includes some geographical filter values but instead there are filter values like chicken, eggs, etc..»

For this reason, the filter name was changed. Also, the second filter and the titles of the visualization on the Risk Communication Dashboard had a bad naming in the opinion of several users. This also led to a name change in both of them. User 3 annotated:

«The dashboard includes a lot of information which is not recognizable at first glance.»

To counteract this, the renaming of the filters and titles as well as the extension of the tooltips

should help. In addition, the visualizations were connected to allow mutual filtering by clicking on a datapoint. This gives the user more flexibility to explore all information contained in the dashboard at once. This feature was then later praised by user 6 in the second iteration.

Summarized it can be said that 91.67% of the negative comments occurred in the first iteration of the usability testing. In the second iteration some of the new features and designs were commented in a positive way by the participating users.

5.2.3 SUS Questionnaire

After completing all tasks related to a given dashboard, the test users had been asked to complete the ten questions of the SUS questionnaire (Brooke, 1996) for the respective dashboard.

To assess the overall level of usability and the degree to which the dashboard improved from iteration 1 to iteration 2, the SUS scores were analyzed. A linear mixed model was estimated with dashboard (Relative Risk Dashboard, Contribution to Risk Dashboard, Impact Potential Dashboard, Risk Communication Dashboard) and testing phase (iteration 1, iteration 2) as fixed factors and user ID (users 1 to 8) as a random factor. The model was estimated by restricted maximum likelihood and showed a moderate fit ($R^2 = .16$, adjusted $R^2 = .14$, $-2 \ln L = 407.67$).

The fixed-effect tests indicated that there were significant differences between the four dashboards ($F_{3,306} = 3.34$, $p < .05$). Although the average overall SUS score improved for all dashboards from iteration 1 to iteration 2, the variation in the data was so high that the effects of testing phase and the interaction between dashboard and testing phase did not reach significance. The scores are plotted in Figure 15.

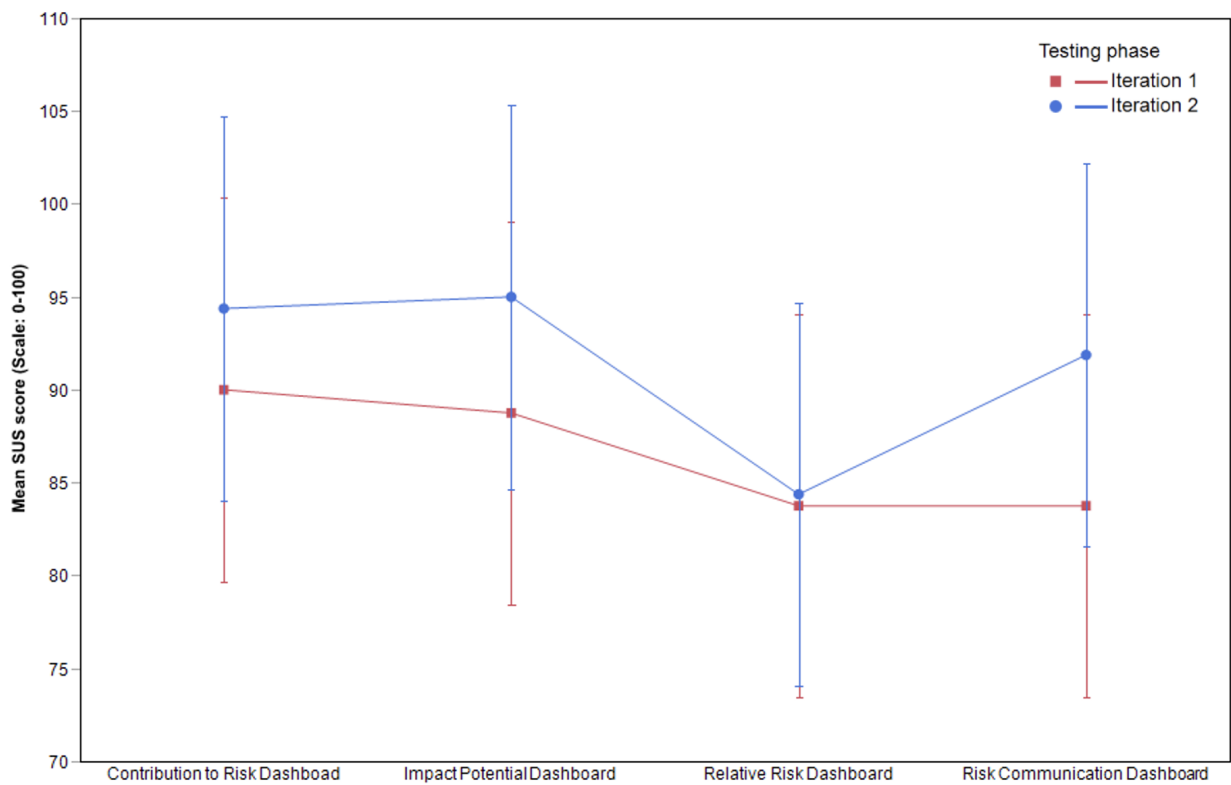


Figure 15: Mean SUS Score as a Function of Dashboard and Testing Phase (Error Bars indicate 95% Confidence Intervals)

Based on the benchmark SUS assessment guidelines by Bangor et al. (2008) the scores obtained by the iteration 2 versions of the four dashboards can be perceived as acceptable and all scores are located in the range from good to excellent. The calculated mean SUS scores for the second iteration are shown in Table 5 and set into the benchmark assessment.

Table 5: Mean System Usability Scale Scores for Dashboards after Testing Iteration 2

Dashboard	Mean SUS Score (Iteration 2)	SD	Benchmark assessment
Relative Risk Dashboard	94.38	17.44	Excellent
Contribution to Risk Dashboard	95.00	15.19	Excellent
Impact Potential Dashboard	84.38	20.17	Good
Risk Communication Dashboard	91.88	17.35	Excellent

Only the Relative Risk Dashboard does not reach the excellent level. However, it is very close to this level, since according to Bangor et al. (2008) a product is classified as excellent from a SUS score of above 85. Since only dashboards with a SUS score below 70 should be revised again before they are published, the SUS Score of 84.38 of the Relative Risk Dashboard is satisfactory, and the dashboard does not need to be revised any further.

6 Discussion

This section is dedicated to the discussion of the results of the two studies mentioned above and to answering the RQs. As a basis for RQ I, the theoretical foundations from Chapter 2.1 as well as the results from the Study 1 in Chapter 4 were used. Whereas RQ II is answered by concerning the literature and previous research from Chapter 2.2 and the findings from Study 2 in Chapter 5. For a better overview, the respective RQ is listed again at the beginning of each paragraph. In addition, the last paragraph gives some suggestions for the design of similar dashboard solutions.

How should dashboards be designed to visualize complex research data and how should they be made accessible to external stakeholders?

Before starting visualizing the research data, it has been showed that it is essential for developers to have a full understanding of the underlying research project's purpose. It was extremely important to understand the research of the SafeConsume project and their findings before starting with the development process of the dashboard. Reading their reports with the most important findings as well as talking to project members, helped a lot to get an overview of the project. If the developer is good enough informed about the project's purpose he can start to look at the raw data. Research data is often collected from different sources and includes some complex measurement units or specific topic related variables. Therefore, the developer must work out a solid data understanding, this includes knowing the origin of the data, how the data was collected as well as the meaning of specific variables, which is also recommended by Few (2006). The underlying data of the present project includes relatively complex epidemiological and statistical material (such as incidence rates, prevalence, odds ratios, relative risks, etc.). These values had to be well known and understood in order to present them in an understandable way to the end users. As David (2021) mentioned if once a good understanding of the project and the data behind it is established, a good basis to start with the dashboard development process is provided.

When designing the dashboards for the research data, it is important that all requirements are precisely defined. According to Chakrabarti et al. (2011) it helps in this step to include some possible end users, like it is done in this work with some project members of the SafeConsume project, which can exactly say which information should be visualized and how it should be visualized. Once all requirements were defined, the ideation phase started. It is recommended by Pillai (2020) to provide some flexibility to this creative step to make the prototype as finished as possible. Therefore, the design- thinking approach is applied as described in Chapter 4.1. This process allows the developer to already incorporate feedback from SafeConsume team members into the prototypes. This ensures that the research data is presented and visualized in an efficient way when it comes to implementation.

There are a lot of design guidelines mentioned in Chapter 2.1.4 which all need to be considered when designing a dashboard. It turned out that the guideline "data needs context" from Few (2006) is really important for designing dashboards with complex research data, since the possible end user can be new to the topics of the SafeConsume project. Therefore, the developer has to find a way to include background information on the dashboard in a nice way that the user can understand and interpret the data. In this work this was done by putting context into different design elements of the dashboard: On every dashboard on top an info icon is placed. When the user hovers over this icon, he can find some more detailed information about the data sources, visualizations and complex terms used in the dashboard. Furthermore, if a user hovers or click on a value in a visualization, the values are set into its context by providing more information about its measurement unit and the meaning in the SafeConsume project. Since the dashboards will be integrated on the website of Nofima, there is also the possibility to insert a text that puts the dashboard in the context of the research work. All these used possibilities to provide background information should make it easier for the user to understand and interpret the complex research data.

Since there are now many different software tools for dashboard development, it is important, as Pillai (2020) also mentioned, to compare the dashboard requirements with the capabilities of the software tools. In this work, firstly Microsoft PowerBi was chosen as the tool for the dashboard development. However, it then turned out that the dashboards can only be made accessible to other PowerBi users and not to the general public. For this reason, the software was changed to the Tableau software which brings benefits like the ease of publishing dashboards and low cost maintenance. This change of software caused a loss of time, since the dashboards had to be developed again from scratch. That it is why it is important to analyze the different tools in advance to prevent such complications later. Furthermore, Tableau offers an online platform, called Tableau public, where users can provide data visualizations to the public for free. They also give research projects the possibility to easily update the dashboards or integrate the dashboards into their website. Interested parties can also download the dashboard to view the raw data behind it or further develop the dashboard by themselves. Therefore, Tableau software can be recommended as tool for make research data available for external stakeholders.

How can high usability be achieved and how can it be evaluated?

The created dashboards should be easy and intuitive to use for the end users. According to Dumas and Redish (1999) conducting usability tests with possible end users should help to achieve high usability. Therefore, for every created dashboard some tasks that should be solved and a SUS questionnaire where given. The performance of the users were measured in time and completeness of received tasks and by calculating the SUS score from the filled out survey. This is how most of the studies evaluate their dashboards in terms of usability and can thus be considered as a commonly used procedure to evaluate the usability of dashboards (Hamm, Money, Atwal, Ghinea, 2019; Martinez et. al, 2018; Nitsch et al., 2019; Schall et al., 2017; van Lieshout et al., 2016; Wu et al., 2019). Since the focus in this evaluation was on the efficiency and effectiveness in working with the dashboard, the tasks need to cover all functionalities and visualizations of the dashboards.

The results show, that both in term of task performance and SUS score achievement, the first dashboard about the Relative Risk performs the worst. It reached the lowest task completeness as well as the lowest SUS score. This is probably related to the fixed sequence of dashboards occurring in the procedure of this study, meaning that the Relative Risk Dashboard was always the first evaluated dashboard by the users. Maybe it would have been better if the order of the tested dashboards were randomized during dashboard testing. On the other hand, it also speaks for the consistency in design and structure of the various dashboards, because once the end user has become accustomed to the first presented dashboard, he could deal with the subsequent dashboards very easily and intuitively. It can be said that this consistency in design and structure of the dashboards has a remarkable influence on usability. Therefore, this consistency in design and structure should be considered when creating multiple dashboards for the same end user group. However, it would be of interest to further investigate whether an other dashboard would be more intuitive and would thus pose a faster handling process of the other dashboards.

Good layout and clear naming of the titles make navigation easier for the user and therefore increase usability. But an exciting finding which is not discussed in many studies is the used color schema for the dashboard which was found to have an influence on the usability of the dashboards. Calzon (2021) best case it to use one color with several shaded levels for the whole dashboard. Few (2006) also recommend to use as few colors as possible and always adapt them to the corporate identity of the dashboard publisher. This ensures that the dashboards are attractive and that you can recognize the publisher of the dashboards by the design. However, it has been shown that this can also lead to problems, since it is more difficult for the end user to distinguish the colors. Thus, he cannot interpret the visualizations as quickly as when various colors are used, which had been shown in the results of the task analysis in Chapter 5.2.1. Therefore, this was adapted and other colors than the logo of Nofima were used for the color code of the visualizations to reach higher usability. In the second iteration of the usability testing, this color change was assessed as positive by the users leading to an increased usability of the dashboards. It is recommended to determine the choice of colors thoroughly and also to test them with end users and not only listen

to the opinion of experts.

During solving the tasks, the participants were asked to think aloud. Most of the testing participants commented about things they don't like about the design or when they were confused of how to complete a certain task. The specific points they criticized about are listed in Chapter 5.2 in more detail. Moreover, it was noticed that the users rarely came up with a suggestion for improvement themselves. This is probably due to the fact that they are not familiar with the field of data visualization and are thus not used to bring in new ideas or suggestions about improvement possibilities such as new ideas for chart types or additional features. All users have analyzed the dashboards in a relatively similar manner except the one testing participant with experience in dashboard design. Compared to the other users, this participant also gave some interesting ideas to improve the usability of the dashboards. In a further iteration or for future testings it would make sense to involve more experts from the field in order to gain such inputs. Furthermore, it would make sense to also include the heuristic evaluation method mentioned in Chapter 2.2.2. This has not been done in this study as the focus was on user feedback. Dowding et al. (2019) stated that if the focus is on user feedback rather than expert feedback no heuristic evaluation is necessary. When more experts participate in the usability testing for the reasons stated above, it would certainly make sense to test the dashboards according to Nielsen's(2020) ten heuristics as well.

Implications for the design of similar dashboard solutions

The results found in this study are based on dashboards of epidemiological and food related research data. However, most of the insights gained in this thesis can also be used for the development of dashboards with other research data. Independent of the context of the research data, research data always brings the problem of being complex and hard to understand for external stakeholders when no context is added. Therefore, every dashboard for research data need to include this background information somehow in the dashboards. This thesis showed some different ways to do handle this issue. Furthermore, the above discussed findings in proceeding with

the design process and design guidelines, the software selection and the usability evaluation can be interesting to read for dashboard developers in general.

Most dashboards deal with dynamic data, which means that the data continually changes after it is recorded. In this thesis the used data are static, which means that the data set remains the same after it is collected. In principle, the solution developed here could be easily adapted to a situation where the data are dynamic, for example if it was decided that the studies and surveys of the SafeConsume project should be repeated in order to collect new and more up-to-date data. This would necessitate to a regular update of the data underlying the dashboards by the future owner of the dashboards (Nofima). As already described, this is not a big deal, since the used software Tableau offers the possibility to easily update the dashboards through Tableau public. However, the design of the dashboards would be needed to be changed because the user wants to compare the different years. Therefore, a year filter should be added on every dashboard on the top right corner, where you can select a single year or a time period. This time dimension would have to be added in each visualization to see developments in the data over time. Furthermore, some new types of visualization could be added or some existing need to be replaced. Then for giving the user the possibility to explore the changes over time, it makes more sense to include line charts or bar charts than tables. And of course we recommend to test the adapted dashboards with data over time again before they are published.

It would also be exciting to know what dashboards would look like with verbal data and how different the whole design process and usability testing would be. As part of the SafeConsume project, WP1 also collected large amounts of qualitative observation protocol data that has not yet been integrated into the dashboard. It would certainly be possible to create a dashboard for this kind of data as well. An idea would be to create an interactive plot based on correspondence analysis, instead of the word clouds usually seen in dashboard visualizations of content analysis results (Few, 2006). It should be possible to click on a data point in the correspondence analysis display. If done so, one level deeper, the analysis of the topic of the data point would be reached. Each data point

would represent a topic from WP1. It would also be interesting to be able to choose the labels used for the correspondence analysis. This could be determined by the user using a filter. Perhaps this is also of too much leeway for the user and would overtax him, since as well a certain know-how is presupposed. There would be the option to only put choices in the filter that make sense for creating a correspondence analysis. Nevertheless, this should be assessed with interviews with possible end users and then carefully integrated as well as thoroughly tested.

7 Conclusions and Recommendations

In this thesis, interactive dashboards for complex research data were designed, implemented and evaluated. The main focus was on the whole development process of the dashboards started from the requirements gathering until the launch of the dashboards. It was shown how crucial it is for the developer to deeply understand what the research project is about. If once a good understanding of the project and the data behind it is established, a good basis to present them in an understandable way to the end users is given. In this step, it is recommended to include some project members which can recommend which information should be visualized and how it should be visualized. It is important to invest enough time in the described step in order to be able to define exact requirements for the dashboards.

After the requirements were defined, the design phase started. A lot of design guidelines in literature exist which need to be considered when developing a dashboard from scratch. It turned out that the guideline, saying that the data needs context in order to be interpreted correctly by the end users, is really important for designing dashboards with complicated research data. Since the possible end user can be new to the topics of the research project, context information helps to orientate himself and understand the presented data correctly. Therefore, the developer has to find a way to include background information on the dashboard in a nice and understandable way.

Throughout the process, it has also become clear that the choice of software for the implementation of the dashboards is fundamental. Therefore, the characteristics of various tools for dashboard development should be analyzed in advance and then compared to the required specifications of the desired dashboards in order to make the best possible choice. It has become apparent that the Tableau software can be recommended as a tool for making research data available in an appealing and coherent way for external stakeholders. This software guarantee as well low maintenance cost and comfortable update possibilities for the project owners.

Except of creating the dashboards, another major part of this thesis was the conduction of the usability tests with eight participants. These users tested the dashboards in terms of functionality and usability. Therefore, for every created dashboard some tasks that needed to be solved and a SUS questionnaire needed to be filled out where given. The performance of the users was measured in time required to fulfill the tasks and the achieved completeness of the tasks and by calculating the SUS score from the filled out survey. With the testing the weaknesses were pointed out and the dashboards were revised and brought to a level that they can be published. The results of the testing show that the dashboards are of good quality and fulfill their functionalities.

The main outcomes of this thesis are the dashboards, which are handed over to the research institute Nofima after the completion of this work. They will integrate them on their website and make them accessible for all interested parties. Furthermore, some more data sets of the underlying research project need to be visualized in form of a dashboard. I am honored to be asked to further work for the SafeConsume project after submitting my thesis. I am looking forward to continue creating dashboards and contributing to other tasks and thus helping bringing this research project to an end.

8 Bibliography

Albert, B., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Waltham: Morgan Kaufmann.

Asif. (2016). 10 Remarkable Principles To Design Informative Dashboards. Retrieved March 21, 2022, from <http://blogs.quovantis.com/ten-remarkable-principles-to-design-informative-dashboards/>

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human-Computer Interaction*, 24(6), 574-594. <https://doi.org/10.1080/10447310802205776>

Barnum, C. M. (2010). *Usability testing essentials*. Burlington: Morgan Kaufmann.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, A. L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 189–194). London: Taylor Francis.

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, 8(2), 29–40.

Calzon. (2021). 23 Dashboard Design Principles Best Practices To Enhance Your Data Analysis. Retrieved March 21, 2022, from <https://www.datapine.com/blog/dashboard-design-principles-and-best-practices/>

Chakrabarti, A., Shea, K., Stone, R., Cagan, J., Campbell, M., Hernandez, N. V., & Wood, K. L. (2011). Computer-based design synthesis research: an overview. *Journal of Computing and Information Science in Engineering*, 11(2). <https://doi.org/10.1115/1.3593409>

Cordis. (2022). *SafeConsume: Safer food through changed consumer behavior: Effective tools and products, communication strategies, education and a food safety policy reducing health burden from foodborne illnesses*. Retrieved March 20, 2022, from <https://cordis.europa.eu/project/id/727580>

David, M. (2021). Dashboard Design Process. Retrieved March 20, 2022, from <https://dataschool.com/how-to-design-a-dashboard/dashboard-design-process/#:~:text=The%20dashboard%20design%20process%20starts,%2C%20get%20feedback%2C%20and%20iterate>

Djuraskovic, O. (2022). Big Data Statistics 2022: How Much Data is in The World?. Retrieved April 13, 2022, from <https://firstsiteguide.com/big-data-stats/>

Dover, C. (2004). How dashboards can change your culture. *Strategic Finance*, 86(4), 42.

Dowding, D., & Merrill, J. A. (2018). The development of heuristics for evaluation of dashboard visualizations. *Applied clinical informatics*, 9(03), 511-518. <https://doi.org/10.1055/s-0038-1666842>

Dowding, D., Merrill, J. A., Onorato, N., Barrón, Y., Rosati, R. J., & Russell, D. (2018). The impact of home care nurses' numeracy and graph literacy on comprehension of visual display information: implications for dashboard design. *Journal of the American Medical Informatics Association*, 25(2), 175-182. <https://doi.org/10.1093/jamia/ocx042>

Dowding, D., Merrill, J. A., Barrón, Y., Onorato, N., Jonas, K., & Russell, D. (2019). Usability evaluation of a dashboard for home care nurses. *Computers, informatics, nursing: CIN*, 37(1), 11. <https://doi.org/10.1097%2FCIN.0000000000000484>

Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*. Norwood: Ablex.

Eckerson, W. W. (2010). *Performance dashboards: measuring, monitoring, and managing your business*. New jersey: John Wiley Sons.

Febretti, A., Sousa, V. E. C., Lopez, K. D., Yao, Y., Johnson, A., Keenan, G. M., & Wilkie, D. J. (2014). One size doesn't fit all: The efficiency of graphical, numerical and textual clinical decision support for nurses. *Proceedings of the IEEE VIS Workshop on Visualization of Electronic Health Records*.

Few, S. (2006). *Information dashboard design: The effective visual communication of data (Vol. 2)*. Sebastopol: O'reilly.

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A. S., Broschinski, S., & Politi, M. C. (2012). Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology, 31*(3), 286. <https://doi.org/10.1037/a0024850>

Galesic, M., Gigerenzer, G., & Straubinger, N. (2009). Natural frequencies help older adults and people with low numeracy to evaluate medical screening tests. *Medical Decision Making, 29*(3), 368-371. <https://doi.org/10.1177/0272989X08329463>

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22*(5), 392-399. <https://doi.org/10.1177/0963721413491570>

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. *Bmj, 327*(7417), 741-744. <https://doi.org/10.1136/bmj.327.7417.741>

Hamm, J., Money, A. G., Atwal, A., & Ghinea, G. (2019). Mobile three-dimensional visualisation technologies for clinician-led fall prevention assessments. *Health informatics journal, 25*(3), 788-810. <https://doi.org/10.1177/1460458217723170>

Indicative. (n.d.). What Is A Zettabyte?. Retrieved April 13, 2022, from <https://www.indicative.com/resource/zettabyte/>

ISO. (2018). Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts. Retrieved March 26, 2022, from <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>

Jordan, P. W. (2020). *An introduction to usability*. London: CRC Press. <https://doi.org/10.1201/9781003062769>

Kuniavsky, M. (2003). *Observing the user experience: a practitioner's guide to user research*. San Fransisco: Morgan Kaufmann.

Lagha, R. R., Burningham, Z., Sauer, B. C., Leng, J., Peters, C., Huynh, T., ... & Kramer, B. J. (2020). Usability testing a potentially inappropriate medication dashboard: A core component of the dashboard development process. *Applied Clinical Informatics, 11*(04), 528-534. <https://doi.org/10.1055/s-0040-1714693>

Lavery, D., Cockton, G., & Atkinson, M. P. (1997). Comparison of evaluation methods using structured usability problem reports. *Behaviour Information Technology*, 16(4-5), 246-266. <https://doi.org/10.1080/014492997119824>

Martinez, W., Threatt, A. L., Rosenbloom, S. T., Wallston, K. A., Hickson, G. B., & Elasy, T. A. (2018). A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. *JMIR human factors*, 5(3), e9569. <https://doi.org/10.2196/humanfactors.9569>

Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414).

Nielsen, J. (2020). 10 Usability Heuristics for User Interface. Retrieved March 28, 2022, from <https://www.nngroup.com/articles/ten-usability-heuristics/>

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213).

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256).

Nitsch, M., Adamcik, T., Kuso, S., Zeiler, M., & Waldherr, K. (2019). Usability and engagement evaluation of an unguided online program for promoting a healthy lifestyle and reducing the risk for eating disorders and obesity in the school setting. *trients*, 11(4), 713. <https://doi.org/10.3390/nu11040713>

Nofima. (n.d.) About. Retrieved March 20, 2022, from <https://nofima.com/about/>

Pillai, K. (2020). 7-Step Process to Build a Dashboard. Retrieved March 26, 2022, from <https://www.linkedin.com/pulse/7-step-process-build-dashboard-krishnadev-pillai>

Romanelli, S. (2020). Engaging Your Dashboard Users Trough User Testing. Retrieved March 24, 2022, from <https://medium.com/nightingale/engaging-your-dashboard-users-through-user-testing-df786d5afc65>

Rubin, J., & Chisnell, D. (2008). How to plan, design, and conduct effective tests. *Handbook of usability testing*, 17(2), 348.

SafeConsume. (2022). *The Project* Retrieved March 20, 2022, from <https://safeconsume.eu/about/the-project>

SafeConsume. (2022). *Hazards* Retrieved March 20, 2022, from <https://safeconsume.eu/hazards/campylobacter>

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Cambridge: Morgan Kaufmann.

Schall Jr, M. C., Cullen, L., Pennathur, P., Chen, H., Burrell, K., & Matthews, G. (2017). Usability evaluation and implementation of a health information technology dashboard of evidence-based quality indicators. *CIN: Computers, Informatics, Nursing*, 35(6), 281-288. <https://doi.org/10.1097/cin.0000000000000325>

Schooley, B., Hilton, B., Abed, Y., Lee, Y., & Horan, T. (2011). Process improvement and consumer-oriented design of an inter-organizational information system for emergency medical response. *Proceedings of the 44th Hawaii International Conference on System Sciences*, 1–10. <https://doi.org/10.1109/HICSS.2011.351>

SERI. (2021). EU Framework Programmes for Research and Innovation Retrieved April 13, 2022, from <https://www.sbfi.admin.ch/sbfi/en/home/research-and-innovation/international-cooperation-r-and-i/eu-framework-programmes-for-research.html/>

Srinivasan, A., Drucker, S. M., Endert, A., & Stasko, J. (2018). Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE transactions on visualization and computer graphics*, 25(1), 672-681. <https://doi.org/10.1109/TVCG.2018.2865145>

van Lieshout, R., Pisters, M. F., Vanwanseele, B., de Bie, R. A., Wouters, E. J., & Stukstette, M. J. (2016). Biofeedback in partial weight bearing: usability of two different devices from a patient's and physical therapist's perspective. *PLoS One*, 11(10), e0165199. <https://doi.org/10.1371/journal.pone.0165199>

Wexler, S., Shaffer, J., & Cotgreave, A. (2017). *The big book of dashboards: visualizing your data using real-world business scenarios*. New Jersey: John Wiley Sons.

Wu, D. T., Vennemeyer, S., Brown, K., Revalee, J., Murdock, P., Salomone, S., ... & Hanke, S. P. (2019).

Usability testing of an interactive dashboard for surgical quality improvement in a large congenital heart center. *Applied Clinical Informatics*, 10(05), 859-869. <https://doi.org/10.1055/s-0039-1698466>

Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41-59. <https://doi.org/10.1016/j.accinf.2011.08.002>

Zarnegar, G. (2019). Design Thinking Process. Retrieved March 24, 2022 from https://www.pngitem.com/middle/xmoxxm_design-thinking-process-hd-png-download/

9 Appendix

Appendix A - Links to Published Dashboards

Relative Risk Dashboard:

<https://public.tableau.com/app/profile/tim.gr5829/viz/SurveydataDashboard/relativeriskdashboard>

Contribution to Risk Dashboard:

https://public.tableau.com/app/profile/tim.gr5829/viz/imp_by_scope_stage/imp_by_scope_stage

Impact Potential Dashboard:

<https://public.tableau.com/app/profile/tim.gr5829/viz/IPADashboard/IPADashboard>

Risk Communication Dashboard:

<https://public.tableau.com/app/profile/tim.gr5829/viz/Channelsandapproaches/Channelsandapproaches>

Appendix B - Dashboard Usability Test Participants List

Table 6: Usability Test Participants List

User ID	Inclusion Criteria	Familiar with Safe Consume	Testing Iteration
User 1	SafeConsume project member	yes	1.
User 2	familiar with communicating research data	no	1.
User 3	familiar with dashboard design	no	1.
User 4	unknown user	no	1.
User 5	SafeConsume project member	yes	2.
User 6	familiar with communicating research data	no	2.
User 7	SafeConsume project member	yes	2.
User 8	familiar with communicating research data	no	2.

Appendix C - List of Tasks for Dashboard Usability Testing

Relative Risk Dashboard:

- Task 1: Which vulnerable group has the highest healthcare demand? And how much of it is attributable to behavioural risk?
- Task 2: Which greater region has the highest productivity loss (workdays lost per household) attributable to behavioural factors? And what is the observed workdays lost in total for this region?

Contribution to Risk Dashboard:

- Task 1: Which stage of the purchase-to-consumption process has the highest contribution to the total behavioural risk? And what type of food stands out?
- Task 2: Where in the purchase-to-consumption process for chicken and eggs do the main problems occur?

Impact Potential Dashboard:

- Task 1: Consider the intervention "Store eggs correctly if using raw": What is this intervention about? Which vulnerable group is most affected by it (in %)?
- Task 2: Consider the intervention "Cook food contaminated by chicken": In which country would this intervention yield the highest DALY reduction in %? Which country has the lowest impact?

Risk Communication Dashboard:

- Task 1: What is the meaning of the lowest row (Average % of countries providing the content) of the overview table? Try to explain it.
- Task 2: Does Switzerland provide any risk communication about the storage of chicken?

Appendix D - SUS Questionnaire

1. I think that I would like to use this dashboard frequently.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

2. I found the dashboard unnecessarily complex.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

3. I thought the dashboard was easy to use.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

4. I think that I would need the support of a technical person to be able to use this dashboard.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

5. I found the various functions in this dashboard were well integrated.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

6. I thought there was too much inconsistency in this dashboard.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

7. I would imagine that most people would learn to use this dashboard very quickly.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

8. I found the dashboard very cumbersome to use.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

9. I felt very confident using the dashboard.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

10. I needed to learn a lot of things before I could get going with this dashboard.

1 2 3 4 5

strongly disagree ☐ ☐ ☐ ☐ ☐ strongly agree

Figure 16: SUS Questionnaire for Dashboard Testing

Appendix E - Test Environment Website for Usability Testing

The test environment website can be accessed through this link:

<https://tmgrmm.github.io/task-based-usability-testing-of-dashboard-prototypes/>

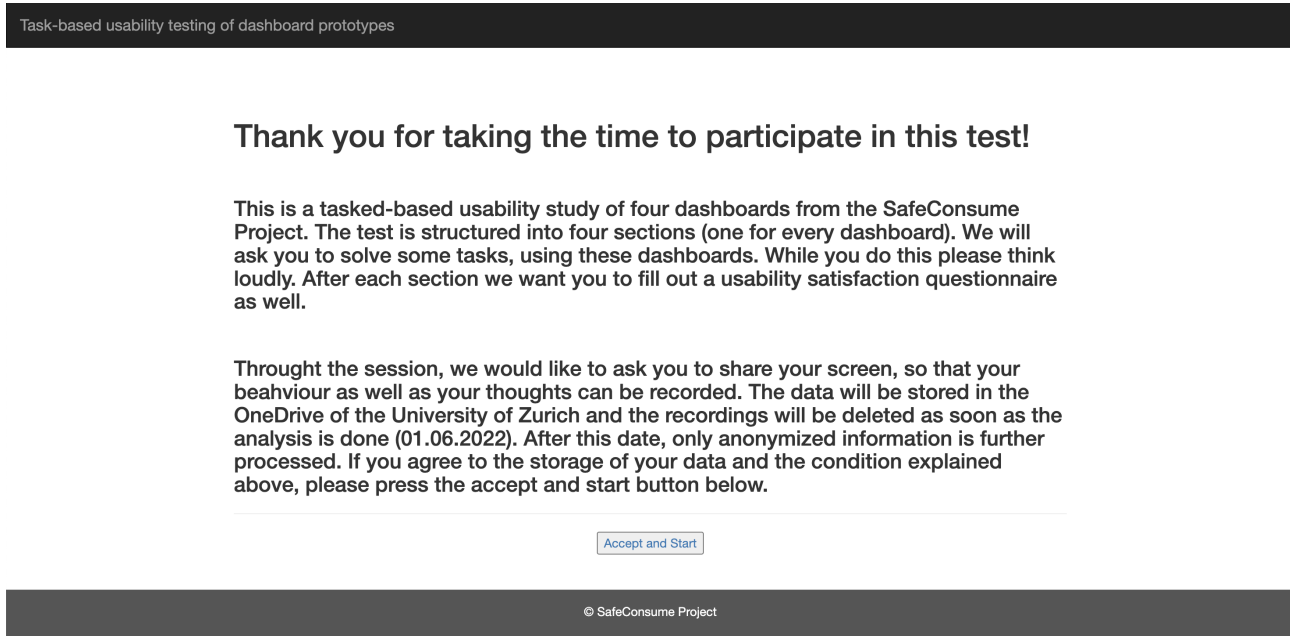


Figure 17: Test Environment Website: Starting Page for Usability Testing