Understanding when customers leave
Defining customer health and how it correlates with software usage

Robert Åman
Abstract

Understanding when customers leave

Robert Åman

More and more businesses today focus on building long-term customer relationships with the objective to secure recurring revenues in competitive markets. As a result, management philosophies such as Customer Success have emerged, which underlines the importance of knowing your customers in order to make them stay. A common way of tracking the well-being of a firm’s customers is the use of customer health scores. Such tools monitor assembled data and indicate whether a customer is doing fine, or is in the risk zone of ending the business relationship. However, there exists little to no consensus on what customer health actually means, or how to distinguish suitable parameters for measuring this concept. Therefore, the purpose of this thesis has been:

To extend the existing knowledge of the business concept customer health, and show how to identify relevant parameters for measuring customer health.

To reach this purpose, a study has been conducted at a software-as-a-service company operating in the field of digital marketing; where methods such as semi-structured interviews, ethnography, web survey, data mining execution and statistical analysis have been used. The results show that software usage differs between active and former customers, with the general tendency that a high software usage indicates a higher propensity to stay as a customer. The study concludes that customer health is best defined as “the perceived value a customer experiences when using a product”. In addition, the parameters that were found to best indicate customer health at the company studied were linked to customers’ software usage as well as their marketing set-up.

Keywords: Customer Success, Customer Relationship Management, customer health, perceived value, satisfaction, loyalty, retention, churn, SaaS
Preface

This report is the final part of the Master Programme in Industrial Management and Innovation and concludes in that sense my Master’s studies at Uppsala University. The study has been conducted during the spring semester of 2017 at the software-as-a-service company Funnel AB. The thesis work has been interesting and enjoyable, and I would like to thank everyone who has been part of the research process.

I am extremely grateful for all the help and feedback that I have received from Sven Hamberg, chief product officer at Funnel AB, and supervisor Göran Lindström, senior lecturer at Uppsala University. Your expertise has been invaluable.

Robert Åman
Uppsala, 2016-07-20
**ANOVA** - **Analysis of variance**: Statistical method used to analyze variance, from which inferences about means can be drawn.

**API** - **Application programming interface**: A protocol for how to programmatically communicate with a system. An API dictates what information can be sent to, as well as retrieved from, the system.

**B2B** - **Business to business**: Commercial activities between two or more businesses, in contrast to business to consumer.

**B2C** - **Business to consumer**: Commercial activities between a business and the final consumer.

**CAC** - **Customer acquisition cost**: The costs associated to acquiring a new customer.

**CES** - **Customer effort score**: Score used by businesses to measure customers’ perception of how good the company is at handling requests and issues.

**CLV/LTV** - **Customer lifetime value**: Estimate of the net profit attributed to the whole future relationship with a customer.

**CRC** - **Customer retention cost**: The costs associated to retaining an existing customer.

**CRM** - **Customer relationship management**: Business approach including practices, strategies and technologies for managing a company’s current and future customer interactions.

**CSAT** - **Customer satisfaction score**: Score used by businesses to measure customers’ satisfaction levels.

**CSM** - **Customer success management**: Business approach that integrates functions and activities from marketing, sales, support, training and professional services, in order to meet the needs of subscription based business model companies. The main idea is that proven value for the customer means value for the company.

**MRR** - **Monthly recurring revenue**: The revenue originating from a company’s monthly subscriptions. Often seen as the most important business metric to track in the SaaS-field. Used at explaining recurring revenue at both company- and customer-level.

**NPS** - **Net promoter score**: Score used by businesses to measure customers’ level of loyalty. The responses are grouped into three groups: Promoters, Passives, and Detractors, and the final score is calculated as the percentage of Promoters subtracted by the percentage of Detractors.

**RFM** - **Recency, frequency and monetary values**: Customer-based value metrics that can be used to segment customers according to their purchase behaviors. Often used in predictive churn models.
**ROAS** - **Return on advertising spend**: Key performance indicator used to determine marketing performance. Shows gross revenue for every dollar spent and is calculated by dividing revenue from ad campaign by the campaign’s costs.

**ROI** - **Return on investment**: Profitability ratio explaining the return on an investment relative its cost. Calculated by dividing the benefit of the investment by the cost of the investment. Expressed as a percentage or ratio.

**SaaS** - **Software-as-a-service**: Software licensing and distribution model in which software is centrally hosted and accessed over the Internet. The software is licensed on a subscription basis.

**SPC** - **Satisfaction loyalty profit chain**: Concept describing the connection between the business-related phenomenon customer satisfaction and loyalty, and their respective influence on customer retention and company performance.

**SQL** - **Structured query language**: Standard programming language for managing, storing and retrieving data in databases.
Table of Contents

1 Introduction ......................................................................................................................... 1
   1.1 Background .................................................................................................................... 1
   1.2 Industry description ....................................................................................................... 2
   1.3 Company description ..................................................................................................... 4
   1.4 Problematization .......................................................................................................... 5
   1.5 Purpose and research questions .................................................................................... 5
   1.6 Limitations .................................................................................................................... 6

2 Theory and literature ............................................................................................................ 8
   2.1 Defining customer health ............................................................................................... 8
   2.2 Customer relationship management ............................................................................. 9
   2.3 Satisfaction-loyalty-profit chain .................................................................................. 10
      2.3.1 Perceived value and its linkage to satisfaction and loyalty .................................... 11
      2.3.2 Customer satisfaction and profit .......................................................................... 12
      2.3.3 Customer satisfaction and retention ..................................................................... 12
      2.3.4 Customer loyalty and profit ............................................................................... 13
      2.3.5 Measuring customer satisfaction and loyalty ..................................................... 15
   2.4 Customer defection and churn prediction models ....................................................... 16
   2.5 Predictive parameters used in previous studies ........................................................... 17

3 Method and execution ......................................................................................................... 19
   3.1 Study design and strategy ............................................................................................. 19
   3.2 Choice of methods ........................................................................................................ 20
   3.3 Interviews .................................................................................................................... 20
      3.3.1 Semi-structured interviews with Funnel’s employees ........................................ 21
      3.3.2 Semi-structured interviews with Funnel’s customers ......................................... 21
   3.4 Ethnography ................................................................................................................ 22
   3.5 Customer web survey .................................................................................................. 22
   3.6 Data mining execution .................................................................................................. 23
   3.7 Analysis of qualitative content ..................................................................................... 24
   3.8 Analysis of quantitative content .................................................................................. 25
   3.9 Critical evaluation of methods .................................................................................... 26
      3.9.1 Interviews ............................................................................................................ 26
      3.9.2 Customer web survey ......................................................................................... 26
      3.9.3 Data mining execution ....................................................................................... 27
      3.9.4 Analysis of quantitative content ........................................................................ 27
   3.10 Criteria in research ..................................................................................................... 27

4 Ethics ................................................................................................................................ 29
   4.1 Ethical principles ......................................................................................................... 29
   4.2 Further ethical considerations ...................................................................................... 30
   4.3 Wider ethical implications ......................................................................................... 30

5 Empirical results ................................................................................................................ 32
   5.1 Results from employee interviews ............................................................................... 32
   5.2 Results from customer interviews ............................................................................. 33
   5.3 Results from customer web survey ........................................................................... 34
5.3.1 Quantitative results from customer web survey

5.3.2 Qualitative results from customer web survey

5.4 Results from data mining execution

6 Analysis

6.1 Customer health and the relationship between satisfaction, loyalty and retention

6.2 Parameters indicating customer health

6.2.1 Likelihood to stay and NPS versus data mining parameters

6.2.2 Status-variable versus data mining parameters

6.3 Critical evaluation of results and analysis

7 Discussion and conclusions

7.1 Theoretical implications

7.2 Practical implications

7.3 Conclusions

7.4 Further research suggestions

8 References

9 Appendices

9.1 Appendix A: Interview template for interviews with Funnel’s employees

9.2 Appendix B: Interview template for interviews with Funnel’s customers

9.3 Appendix C: Web survey- Six short questions about Funnel

9.4 Appendix D: Distribution of quantitative survey answers and the re-grouping of variables

9.5 Appendix E: Qualitative answers from web survey and their clusters

9.6 Appendix F: Test of linearity
List of figures

Figure A: The satisfaction-loyalty-profit chain .................................................................11
Figure B: The satisfaction-retention link .............................................................................13
Figure C: The affective and cognitive dimensions of loyalty ...............................................14
Figure D: The data-mining process ......................................................................................16
Figure E: The distribution of responses for the support-question in the survey ..................35
Figure F: The distribution of the Likelihood to stay-variable between the NPS-segments .....36

List of tables

Table A: Crosstabulation with the NPS-score as row variable and Likelihood to stay as column variable .............................................................................................................36
Table B: Chi-square test for the crosstabulation in table A .................................................36
Table C: The extracted and examined data mining parameters ............................................37
Table D: ANOVA-test with Likelihood to stay as factor and the data mining parameters as dependent variables ..................................................................................................................39
Table E: Test of significance for the ANOVA-test in table D .............................................40
Table F: Linear stepwise regression analysis with Likelihood to stay as dependent variable and the data mining parameters as independent variables .........................................................41
Table G: ANOVA-test with Status as factor and the data mining parameters as dependent variables .................................................................................................................................43
Table H: Test of significance for the ANOVA-test in table G .............................................44
Table I: Binary logistic regression analysis with Status as dependent variable and the data mining parameters as independent variables ........................................................................................................45
Table J: Classification table for the binary logistic regression analysis in table I ...............45
Table K: The null model for the binary logistic regression analysis in table I .....................45

List of formulas

Formula A: Linear regression equation with Likelihood to stay as dependent variable and Activity ratio as independent variable ........................................................................................................40
Formula B: Logistic regression equation with Status as dependent variable and Activity ratio and Number of source types as independent variables ..................................................44
1 Introduction

This chapter will present the background, purpose, and limitations linked to this Master’s thesis project that has been conducted on behalf of the software company Funnel AB. The text will start with a brief background of the problem area and continue to describe the company and their expressed mission for this Master’s thesis work. Thereafter the purpose and research questions will be specified as well as what limitations that will affect the execution of this study. The chapter ends by describing the disposition for the rest of the report.

1.1 Background

Loss of customers has always been a central problem in the field of business. Today, however, a lot of companies experience an increased concurrence due to globalization and greater means of comparison-shopping through the web, resulting in an even harsher business climate where each customer relationship is of importance for the firm (Waxer, 2011). This has led to more and more companies changing their main efforts from acquiring new customers to working more actively in retaining their existing ones, in that way shifting their focus towards more long-term relationships (Tamaddoni Jahromi, Stakhovych & Ewing, 2014). This is especially important for firms working in the business to business (B2B)-segment, since B2B-companies often have fewer customers to tend to, making each customer relationship more valuable and economically important to the firm (Stevens, 2005).

Metrics such as customer lifetime value (CLV) has emerged as a way to estimate what economical contribution each customer will have to a business. There are plenty of ways to calculate this metric, varying from vary basic mathematical formulas to advanced predictions of future purchase patterns through the use of sophisticated statistical models. The outcome is although always an estimated value that indicates how profitable each customer will be over his or her lifecycle. This value is often looked at in relation to two other metrics: customer acquisition cost (CAC) and customer retention cost (CRC). The metrics describe the costs associated with acquiring a new customer and the costs associated with retaining that customer. These three metrics result in a basic mathematical equation that each business must keep an eye on in order to survive in the long run. In other words, if the CAC and CRC exceed the estimated CLV, the customer is not worth acquiring in the first place. Although the importance of these metrics might seem vital, far from all companies measure and follow up on them. In fact, according to Khalid Saleh at Invesp (2015), 76% of companies view CLV as an important metric to track in their organizations, yet only 42% are able to calculate it correctly.

Regarding costs associated with acquiring and retaining customers, studies have shown that acquiring a new customer is between five to 25 times more expensive than retaining an existing one (Gallo, 2014). In addition, Frederick Reichheld and Phil Schefter (2000) note that by increasing retention rates by 5%, profits have the potential to go up somewhere between 25% to 95%. Consequently, retention strategies generally have a higher return on investment (ROI) than strategies focused on acquisition, leading to the general belief that a company should focus most marketing resources on keeping their existing customers (Tamaddoni
Jahromi, Stakhovych & Ewing, 2014). With this in mind it is surprising to see that only 18% of companies put the greater focus on retaining customers versus 44% that instead emphasize customer acquisitions, especially since existing customers also have a higher tendency to spend more money and try out new products compared to new customers (Saleh, 2015). In other words, companies can yield great rewards by focusing more on retaining and handling their existing customers, instead of using all their efforts in acquiring new ones.

Fundamental to all retention strategies is the challenge to minimize customer churn. Churn is defined as “the tendency for customers to defect or cease business with a company” (Kamakura et al., 2005) and is in other words the opposite of customer retention. Put differently, if a company’s churn rate is higher than its acquisition rate, the company will experience a continuous loss of customers, eventually leading to the drainage of the whole customer base. Such a development implies the end of a business. Because of that, it makes perfect sense to investigate when and why customers churn when focusing on customer retention.

Even though this logic might seem straightforward, there are great difficulties linked to getting this kind of insight. To begin with, the reasons to why customers churn can be numerous and is often revealed too late for a company to act upon, if ever. Additionally, this information is hard to get by other means than directly asking the customer in question, making it a type of information that is hard to achieve. One way of totally bypassing this problem is to send out incentives to all customers, and in that way reaching those customers that are thinking about ending their business relationship. However, such a strategy will also target customers already intent on staying, and will in that way waste a lot of a company’s resources (Tamaddoni Jahromi, Stakhovych & Ewing, 2014). Additionally, since the company has no possibility to frame the incentive differently after each customer need, the incentive will only be interesting to some customers and therefore only prevent a few potential churners from leaving.

Luckily, innovations in IT have increased the amount of data that can be collected about each customer, and access to this data opens up for more targeted marketing approaches (Burez & Van den Poel, 2007). For e-commerce companies this data can for example include purchase history, while it for companies offering software solutions can include monitored product usage. Therefore, by analyzing these big amounts of data, companies can get insight into how customers behave before they churn and be able to pinpoint potential churners before they actually defect. With that information at hand, a company can then target those customers with customized incentives or offer more direct assistance, which saves both resources and have a higher chance of satisfying customers’ needs. In other words, by understanding when customers leave, companies can start working more pro-actively with their customer base and make their valuable customers stay to a higher degree.

1.2 Industry description
A type of companies that have perfect conditions for assembling and analyzing customer data is software-as-a-service companies. Software-as-a-service, abbreviated SaaS, is a licensing
and delivery model for software in which the vending company hosts the application and guarantee to deliver functionality on demand to its customers. In contrast to traditional software distribution where customers have to buy their own copy of software license and install it, SaaS applications are typically accessible through the Internet using a web browser. Additionally, SaaS companies often bill their customers on a subscription basis (Butterfield & Ngondi, 2016).

The change to subscription-based business models has been on the rise the last couple of years and has led to a new phrase emerging first coined by the software company Zuora, namely “the subscription economy” (Rao, 2014). In such an economy, a customer pays for a product or a service by subscribing to it, instead of paying for it once. The business model gives more flexibility to the customer in terms of pricing options, which in turn might be seen as a way to reduce the economic risk for the customer linked to the purchase. In addition, it has also been proved to increase the intimacy between the firm and their customers (Whitler, 2016). This shift further underlines the importance of managing and helping the existing customers and emphasizes the importance of each customer relationship for the economic performance of the company. In a subscription-based economy a company’s main focus should therefore lie on the customer rather than on the product or transaction (Whitler, 2016).

New business and delivery models tend to not only affect financial parts of a firm but also its organizational work and management. A management philosophy that has evolved out of SaaS and subscription-based business models, is Customer Success, which offers a better way of handling customer relations when working with such business models (Reni, 2015). Before the evolution of subscription software, all software systems were offered as on-premise softwares that were installed and run on the customers’ own computers. The implementation, installation and service of these systems meant great investments for the customers, leading to high-switching costs that made customers stick with the providing company. However, when software is offered on a pay-by-month basis and doesn’t include any difficult installation-process, the decision to switch provider is both easy and cheap. In such a way, the SaaS model gives more power to the customer, and less power to the software-provider, a change that has made the old ways of handling customers less successful (Reni, 2015). The solution to how to engage with your customers in this new environment was named Customer Success and is a perfect example of a management philosophy that stems from the business field, rather than the theoretical field.

The philosophy’s center core is that all parts of a firm should work towards creating as much value as possible for the customer so that he or she reaches his or her goal linked to the product or service usage. The aim is in other words to make each customer as profitable and productive as possible by making sure that the customer gets maximum value out of his or her purchase (Smilansky, 2016). This will in turn lead to recurring profits from retained and satisfied customers, as well as a growing revenue linked to upsells, acquisitions of new customers and referrals from existing ones. The previously mentioned churn-metric is of great importance when working with Customer Success since it can be argued to be a fairly good way of measuring the success of the company’s customer work.
1.3 Company description

The company that has been the object of study in this Master’s thesis work, Funnel, is a SaaS company with headquarters in Stockholm. The company offers a software tool with the same name, which main goal is to ease B2B-customers’ work with compiling and analyzing their advertising data. Roughly described almost every company nowadays works with some kind of online advertising or marketing, most often via several different web channels and platforms. Each advertising platform registers the number of clicks and impressions that relate to a company’s advertising efforts on the site and summarize this together with the cost of advertising. The outcome of the advertising efforts however, such as purchases and sign-ups, is nothing that the advertising platforms have any statistics about. This information can instead be found in the well-known web analytics service Google Analytics.

Consequently, for a company to get insight into how well each web channel or platform performs in terms of key metrics such as return on investment (ROI) and return on advertising spend (ROAS), the advertising company needs to extract data from each advertising platform as well as data from Google Analytics and put this together manually in a separate spreadsheet. Although this can be an easy task for a company that only advertises on a few web platforms, the amount of time related to this procedure quickly increases with the number of ad platforms and websites advertised. Furthermore, due to the time requirement, this is a process that most companies only perform once or twice a month, with the outcome that the data presented is most of the time not up to date. Lastly, such manual labor is also prone to human errors.

What the software tool Funnel does is to automatically gather data from the various advertising platforms, relate that to the client’s Google Analytics-account, and compile everything in order to give a complete picture of a company’s advertising efforts and results. In such a way, Funnel’s customers get a quick and accurate overview over how well their different advertising channels perform in terms of investment and revenue. The advertising data is also continuously updated in the tool to give each customer the means to monitor and reallocate their advertising efforts.

Funnel offers three products: Google Analytics Upload, Dashboards & Reports, and Funnel API. In this study, Dashboards & Reports has been the only product studied. The reason for this is that Dashboards & Reports is the only product that fully monitors and stores customer usage data, data that has been a requirement for the execution of this study. The product itself, is a web application that customers reach by using a web browser, and each customer gets access to their advertising data by logging into their account with username and password. In such a way, customers get instantly access to all their advertising efforts with additional features such as flexible channel grouping of data, customizable dashboards and easy reporting features. In addition, both product usage data and advertising data is monitored and stored in Funnel’s database to be able to give support and offer back-ups whenever an error occurs.
Funnel has integrated the previously named management philosophy Customer Success into their organization and views a minimized customer churn as one of their two most important business goals. The other one is monthly recurring revenue (MRR), which is the total revenue that stems from the recurring monthly subscription fees. These two metrics are in other words tightly linked to each other and are continuously monitored in order to determine the organization’s performance and progress.

1.4 Problematization
As of today, Funnel is in a growing phase and keeps adding new customers to their existing customer base every month. This expansion calls for better means to monitor the “health” of the customers since a growing number of customers makes it harder to monitor each customer individually and decide where to focus the company’s resources. It is therefore of great importance to start working more pro-actively with the customer base to be able to early on identify users that have problems with their product usage. By identifying such customers at an early stage the Customer Success-team can reach out to the struggling individuals and hopefully prevent them from churning. Unfortunately, the means to succeed with such preventive work is today small due to the difficulty in identifying these struggling customers. It is this dilemma that forms the baseline for this study.

A way to gauge a customer’s health is the use of a customer health score. Such a tool’s purpose is to indicate which customers that seem less satisfied with the tool and might be in the risk zone of churning, as well as to indicate what customers that are satisfied and might be the object to cross- or upselling. A customer health score can in other words meet the previously mentioned need to get a good overview of the customer base and also be able to pinpoint which customers that needs more resources.

With this as background, Funnel has acknowledged the need and wish for a way to determine their customers’ well-being and find out what parameters that might be used to track customer health. When identified, these parameters could be assembled and weighted in order to form a customer health score. This tool would give the Customer Success-team a better and more effective way to track and tend to their customers and could possibly indicate what customers that might be the most profitable for the company. In such a way, the company could get better strategic basic data of what customers or customer segments to focus on. In addition, it could also give hints to the Product Development-team in what features the customers like, and which ones that don’t leverage value to the same extent. Another benefit for the rest of the company could be to allow for better insight into the customer base for individuals that don’t work with the customers on a daily basis, such as the board of directors. Finally, it could also be used as a way to gauge the Customer Success-team’s work.

1.5 Purpose and research questions
This study’s focus lies on the preceding process of creating a customer health score, in other words finding relevant parameters for measuring customer health. It should be noted, that even though the study has a case study design that examines a specific company, the process of defining and analyzing customer health parameters should be seen as a more general
attempt to offer a procedure that can be applied at other companies. That being said, the purpose of this Master’s thesis work is:

To extend the existing knowledge of the business concept customer health and show how to identify relevant parameters for measuring customer health.

To reach this goal the SaaS company Funnel and their customers have been the object of study, and the following research questions have been produced to guide this study forward:

**Research question 1:**
What is customer health and how can previous theory and literature help us understand this concept further?

**Research question 2:**
What is seen as good and poor customer health at Funnel and how can this be identified in customers’ product usage?

**Research question 3:**
What value is the product providing its users and how is this related to customer health?

**Research question 4:**
What parameters can be used to indicate customer health for Funnel’s customer base and are some parameters more important to study than others?

These four research questions will be observed throughout the report and answered by applying suitable methods. To begin with, research question one will be studied by consulting previous literature in chapter two, *Theory and literature*. Thereafter, research question two will be observed by interviewing Funnel’s employees, which results will be presented in chapter 5.1, *Results from employee interviews*. Research question three will then be observed in chapter 5.2, *Results from customer interviews*, by interviewing Funnel’s customers, as well as in chapter 5.3.2, *Qualitative results from customer web survey*, by gathering customer responses from a web survey. Lastly, research question four will be touched upon in several parts of the report. First, in chapter two, *Theory and literature*, by looking at what parameters that have been used for predictive purposes in previous research. Second, in chapter 5.1, *Results from employee interviews*, and 5.2, *Results from customer interviews*, by asking Funnel’s employees and customers about parameters suitable to study. And third, in chapter 5.3.1, *Quantitative results from customer web survey*, and chapter 5.4, *Results from data mining execution*, where the chosen parameters, stemming from the interviews, will be analyzed using different statistical methods. Lastly, the results will be summarized and discussed in chapter 6, *Analysis*, where all of the research questions finally will be answered.

**1.6 Limitations**
Since this case study focuses on studying Funnel’s customers and their interactions with the tool, a lot of the limitations that exist are linked to the company studied. Firstly, only customers using the web application Dashboards & Reports will be studied since it is the only product, out of the three offered, which fully monitors and stores customer usage data.
Secondly, usage data older than January 2016 won’t be part of the study. This is due to the fact that both company and product has developed considerably the last couple of years so looking at usage data further back in time would be similar to looking at the use of a totally different product. Finally, even though the creation of a customer health score is the ultimate outcome of a study like this, the scope of this study will be limited to defining what parameters that can be used to measure customer health. The classification and weighting of parameters as well as the procedure of compiling a final health score will be left to future research.
2 Theory and literature

This chapter will examine what theories and previous literature that exist about customer health as well as what adjacent concepts that can be used to understand the concept further. Literature will also be observed in order to find out what parameters that might be of interest to study. This chapter will therefore provide a foundation for answering research question one, and work as a starting point for answering research question four:

Research question 1:

What is customer health and how can previous theory and literature help us understand this concept further?

Research question 4:

What parameters can be used to indicate customer health for Funnel’s customer base and are some parameters more important to study than others?

2.1 Defining customer health

Customer health is a commonly used concept in the field of Customer Success and though the origin is unclear the concept has a clear analogy with personal health. This implies that a good customer health is something desirable, while a poor customer health is something that the company needs to take care of in order to prevent the customer from “dying”. In 2014 the Customer Success-company Gainsight commissioned the American market research company Forrester Consulting to investigate how customer health is measured and tracked in the business field (Forrester Consulting, 2014). 13 US-based SaaS companies were interviewed in order to get ideas and insight into how companies measure and define customer health and the report could conclude the following:

Customer health correlates to the propensity for churn or growth. A good health score is an indication of customer satisfaction. A poor health score is an early warning signal for churn. (Forrester Consulting, 2014)

What can be read from this definition is that the earlier mentioned analogy to personal health holds: when talking about customer health in a subscription-based company death equals churn, while good health equals retention. However, the definition by Forrester Consulting is rather vague and is just one among many, hence an ambiguity of what customer health actually means seems to prevail in the field. Other definitions mention the level of customer loyalty and delivered value as parts of customer health (Raboy, 2013; Abbott, 2017), hence the concept links to a customer’s well-being in some way. However, it is of interest for this study to understand what the concept customer health actually tries to capture and what we are trying to measure. A good start is therefore to examine the abovementioned concepts of perceived value, customer satisfaction, loyalty, and retention and find out how these concepts might relate to a customer’s health.

In addition to defining customer health, this study’s goal is to identify what parameters that might indicate a customer’s well-being, which is specified in research question four. However, no information exists in the academic field regarding customer health scoring and it
is apparent that there exists no definite procedure for identifying these parameters in the business field either. A few software providers offer customer health scoring services, though these softwares often comes with predefined parameters, which the company in question then can choose to monitor, alternatively specify their own. Such parameters can for example be financial data, Customer Relationship Management (CRM) data, product usage data, support tickets, and direct customer feedback (Forrester Consulting, 2014). However, what parameters that are important to study are very much company specific and depends on the product or service provided as well as what possibilities the company has to monitor and track product usage. A lot of this knowledge surely exists within Funnel’s walls, but it can also be of interest to see what previous research say about monitoring customers’ well-being.

The lack of theory on both customer health and how to measure it calls for an investigation of business fields related to Customer Success. The next section will therefore investigate what a somewhat older business field, Customer Relationship Management, can give for insight to this study.

2.2 Customer relationship management

A closely related business field to Customer Success is Customer Relationship Management, abbreviated CRM. This business approach has been around for several years and has changed many organizations and the way they are managed by placing the customer in the heart of the organization and focus on the interaction with current and future customers. Because of different industries great interest in the approach and its continuous development, several definitions of CRM have emerged. This text will follow the strategic level-approach as described by Viba Kumar and Werner Reinartz (2012, p.4), in which CRM is described as a process for how to achieve customer centricity in the marketplace and spread knowledge of the customer to all parts of the organization.

According to Kumar and Reinartz (2012, p.19) the rise and evolution of CRM has been pushed forward by a shift from transactional to more relationship-based markets. This shift has increased the importance of knowing how to interact with your customers. A central part of CRM is therefore to recognize different types of customers in order to develop separate strategies for how to interact with them. The aim of such strategies can for example be to improve relationships with the most profitable customer, find new customers that are likely to be profitable, and develop strategies for less profitable customers that cause the company to lose money. Customer value is in other words a key concept in CRM, which is further underlined in the following definition:

CRM is the strategic process of selecting customers that a firm can most profitably serve and shaping interactions between a company and these customers. The ultimate goal is to optimize the current and future value of customers for the company. (Kumar & Reinartz, 2012, p.5)

This also requires the firm to handle the relations well by delivering expected value and satisfaction, as well as balance the interest of both the organization and the customers. When done right, this can lead to a big competitive advantage for the firm.
Another aspect that has allowed CRM to evolve into such a popular business practice is the data storage technology. The evolution of this technology has led to declined costs in storing data, which in turn has resulted in an exponential growth in data storage capacity (Kumar & Reinartz, 2012, p.20). As a consequence, firms nowadays have the means to collect and analyze big amounts of data about their customers and their transactions. This has opened up for the implementation of CRM-systems that can help the firm get customer insight and make strategic use of the collected data. However, according to Kumar & Reinartz (2012, p.13), too much data can also be challenging, since misapplied analyses often are the outcome when a company is overwhelmed with more data than it can handle.

In the school of mass marketing customers are segmented based on their common needs, which works as input for product development. Products and services are thereafter designed to meet the general needs of these segments. However, thanks to improved information technology and more flexible manufacturing processes, individual customer needs can nowadays be met to a larger extent, something that many firms has turned into a competitive advantage. This has in turn led to a shift from product-based marketing to marketing where the focus is set on the customer (Kumar & Reinartz, 2012, p.19). Such marketing emphasizes the need to retain existing customers, which has led to several firms focusing on increasing the satisfaction levels of their customers. The reason behind this, is the general belief that customer satisfaction leads to retention or loyalty, which in turn leads to increased profit (Kumar & Reinartz, 2012, p.21). This general line of thought is called the satisfaction-loyalty-profit chain and will be thoroughly discussed in the following sections.

2.3 Satisfaction-loyalty-profit chain

The satisfaction-loyalty-profit chain, abbreviated SPC, is a concept that has been around since the beginning of 1990 when several industries started to realize the importance of measuring and handling customer satisfaction (Heskett, Jones, Loveman, Sasser, & Schlesinger, 1994). The underlying idea of the SPC-concept is that improved product and service quality, as well as staff performance, leads to increased customer satisfaction, which in turn leads to increased customer retention. This increased customer retention, which is tightly linked to customer loyalty, will eventually lead to increased firm performance such as increased profitability and can be seen as the ultimate business outcome. Additionally, a lot of research has focused on studying the relationship quality between the firm and the customer, which is believed to further mediate the linkage between satisfaction and retention (Kumar & Petersen, 2012, p.63). See figure A for a graphical illustration of the SPC.
Figure A: The satisfaction-loyalty-profit chain. (SPC) (Own, after Anderson & Mittal, 2000 and Kumar & Petersen, 2012)

Even though this general line of thought might seem intuitive to the reader, empirical findings have only showed mixed support for the SPC-concept (Zeithaml, 2000). Many firms have for example had great problems in converting the conceptual framework into practice (Ittner & Laracker, 2003) while other research studies have had a hard time in validating the linkage between the different sections in the SPC-model (Zeithaml, 2000). For that reason, the different parts of the SPC will be problematized and discussed more closely in subsequent sections, however we will start by examining a concept that is assumed to be the determinant to both satisfaction and loyalty, namely perceived value.

2.3.1 Perceived value and its linkage to satisfaction and loyalty

Perceived value is a concept sprung from equity theory, which examines and describes the relation between the ratio of a consumer’s outcome/input and a service provider’s outcome/input (Oliver & DeSarbo, 1988). In other words, the equity concept describes the process of how a customer evaluates an offering in terms of what is right, fair or deserved in relation to the offering’s perceived cost (Bolton & Lemon, 1999). The customer is assumed to feel equitable treated if the perceived ratio of his or her own outcome/input is similar to the ratio of the company’s outcome/input (Oliver & DeSarbo, 1988). In addition, the customer often measures a company’s outcome/input ratio by comparing the company’s offering with competitors’ (Yang & Peterson, 2004). That being said, perceived value has been found to be related to both customer satisfaction and loyalty. Sirdeshmukh, Singh and Sabol (2002) argue for example that customer value is likely to regulate customer loyalty, and research by Yang and Peterson (2004) found that perceived value is a key driver of customer loyalty. These arguments are further backed up by empirical research from various industries such as telephone services, airline travel, and retail that has recognized perceived value as a major determinant of customer loyalty (Bolton & Drew, 1991; Sirdeshmukh et al., 2002). Customer satisfaction, on the other hand, is often perceived to be a function of perceived service quality (Cronin & Taylor, 1992; Parasuraman, Zeithaml, & Berry, 1988). Furthermore, perceived service quality is linked to the overall perceived value of a product or service, which has shown to have a positive effect on customer satisfaction in several studies (Anderson & Mittal, 2000; Walter, Thilo, & Helfert, 2002). The link between perceived value and satisfaction was also proved to exist in an online service setting (Yang and Peterson, 2004).
The same researchers also pointed on the fact that customer satisfaction might have a mediating role in the relation between customer-perceived value and customer loyalty.

### 2.3.2 Customer satisfaction and profit

A lot of research has focused on studying the assumed direct link between customer satisfaction and profit, in that way bypassing the part of the chain that involves customer loyalty and retention, see figure A. Such a study was conducted by Anderson, Fornell and Lehmann (1994) when they studied a data set from several Swedish industrial companies. Their aim was to find evidence that higher levels of satisfaction leads to greater profits, in addition to looking at the importance of product quality and expectations on satisfaction. The results showed that quality, as well as expectations, correlated positively with customer satisfaction. In addition, the findings supported the authors’ idea that customer satisfaction is a phenomenon that grows stronger with time, which means that short-term variations in quality not necessarily affect the customers’ overall satisfaction. Regarding the financial outcomes, a strong link between return on investment, ROI, and satisfaction was found. In addition, the study suggested that an increased market share might lead to decreased levels of satisfaction, with the explanation that as a company grow it becomes harder to meet individual needs of customers and in that way keep them satisfied. It should be noted that this study was conducted before the breakthrough of CRM, and that CRM approaches might have moderated the risk today. It should also be added that this study’s aim is to avoid such development at Funnel.

Even though the study by Anderson, Fornell and Lehmann (1994), as well as several other studies (Ittner & Laracker, 1998), has shown a positive link between satisfaction and profits, other studies haven’t found any correlation at all (Zeithaml, 2000). An explanation, presented by Kumar and Reinartz (2012, p.26), is that improving satisfaction comes with a cost and if that cost is too high it might take away the positive impacts from profits. The authors also suggest that there might exist optimum satisfaction-levels above which improved satisfaction doesn’t pay off. This was actually the case in a study by Ittner and Larcker (2003) in which a maximum threshold was discovered at an 80% satisfaction-level. In other words, it was found that customers that were 100% satisfied didn’t spend any more money than those satisfied at an 80%-level. However, the costs in bumping customers up from 80% to 100% were substantial.

### 2.3.3 Customer satisfaction and retention

Because of the contradictive results from the research on satisfaction and profit, a lot of studies have instead focused on studying the more closely linked concepts satisfaction and retention, as can be seen in figure A. Research by Zeithaml, Berry and Parasuraman (1996) found for example that satisfied customers are more inclined to hold a stronger repurchase intention as well as recommend the product or service to friends and colleagues. Additionally, research by Bloemer, de Ruyter, and Wetzel (1999) and Oliver (1999) showed that customer satisfaction positively affects customer loyalty. Figure B shows how the two concepts are connected, as supported by several studies (Kumar & Reinartz, 2012). As the figure shows, there exist a nonlinear relationship between the two variables where dissatisfaction has a
stronger effect on retention than satisfaction. The flat part in the middle is often referred to as the zone of indifference (Jones & Sasser, 1995). However, the form of the curve as well as the placement of the sharper corners can differ greatly between industries (Kumar & Reinartz, 2012, p. 27). The context is therefore an important factor when studying this relationship and the linkage strength can be influence by factors such as competitiveness of the market, customers’ switching costs, and perceived risk. Another factor that might affect the link between the two concepts is the measurement chosen for the degree of loyalty, or retention. Mittal and Kamakura (2001) found for example that it is preferred to measure repurchase behavior instead of repurchase intent. A finding that is supported by Trout and Rivkin (2008) who found that 89% of a car brand’s owners said they were very satisfied with their purchase, out of which 69% said they intended to purchase from the same manufacturer again. In reality, less than 20% did so.

Figure B: Illustration of the satisfaction-retention link. The dotted line shows a linear approximation of the nonlinear relationship between satisfaction and retention. (Own, after Anderson & Mittal, 2000)

2.3.4 Customer loyalty and profit
Moving further up the SPC we finally end up at the link between customer retention and profit, see the last sections of figure A. This link has got plenty of attention because of the earlier mentioned high failure rate in linking satisfaction to retention and profit. To begin with it should be noted that while customer loyalty and retention is tightly connected, they aren’t the same thing. Customer retention is a measure of how likely a customer is to come back and repurchase, and is therefore seen as a strong predictor of profitability (Allen, 2004). Customer loyalty, on the other hand, describes the attitudinal state that might evolve as a result from that repurchase behavior (Allen, 2004). Kumar and Reinartz (2012, p. 95) describe it as a “positive emotional or psychological disposition” towards a certain brand. In other words, just because a customer repurchase from a company doesn’t mean he or she is loyal to it. Some factors that can lead to retention but doesn’t imply loyalty are for example inertia, indifference or exit barriers (Reichheld, 2003). Loyalty can further be described as having two dimensions, the affective and the cognitive, as can be seen in figure C.
The affective dimension is the emotional connection that a customer may develop for a certain product, service, or company, and can for example be affected by the relationships with a company’s employees. The cognitive dimension, on the other hand, includes more rational aspects and can include critical assessments of a company’s price, product quality, and problem resolution. That being said, the level of customer loyalty is also assumed to change more slowly, compared to customer satisfaction that may increase or drop due to small things like a minor change in product quality (Allen, 2004).

Rewinding back to retention and financial outcomes, this last part of the SPC-concept has been both hailed and questioned in the field of marketing. The general line of thought is that retention leads to greater profits, since customers that return to do business with a company eventually have purchased enough to cover their own cost of acquisition. In other words, long term customers become more and more profitable over time (Kumar & Reinartz, 2012, p.28). According to Frederick Reichheld (2003) loyal customers are also assumed to:

1. Spend more over time
2. Cost less to serve
3. Generate word-of-mouth to a greater extent
4. Be more willing to pay more for a service or product compared to short-term customers

In addition, they have a higher propensity to focus on more long-term benefits and participate in cooperative activities favoring both customer and firm in comparison to disloyal customers. Such an inclination improves the competitiveness of both actors and reduces the transaction costs (Morgan & Hunt, 1994), which leads to greater profitability. Lastly, retention strategies have shown to lead to a generally higher return on investment (ROI) than acquisition strategies (Lam, Shankar, Erramilli & Murthy, 2004). However, these findings have also been met by doubt from some academics. Kumar and Reinartz (2012, p.29) have for example found limited support for the relationship strength between retention and profit. They mean that in most businesses there exist a segment of loyal customers that aren’t very profitable, and some short-lived customers that generate high profits in a very short amount of time. They argue that loyal customers in fact are often more expensive to serve because they know
their value to the firm and often use this advantage to get premium support or discounts (Kumar & Reinartz, 2002). Such royal treatment quickly eats into profit. Kumar and Reinartz (2012, p.29) conclude that caution should be taken when equating customer retention with profit and that customer value must be taken into consideration when making marketing decisions, not just the level of customer satisfaction and loyalty.

2.3.5 Measuring customer satisfaction and loyalty
Even though it now should be clear that the link between customer satisfaction, loyalty and profit is far from straight and might vary between industries, a lot of companies lack the means to include customer profitability in their customer base analyses, and therefore often use satisfaction and loyalty scores as a proxy measure for customer profitability. A common way to measure such scores is the use of surveys or short questions. Examples of such well-used methods are the Customer Satisfaction Score (CSAT) that indicate the satisfaction-level of the customer, the Net Promoter Score (NPS) that is argued to measure the loyalty-level, and the Customer Effort Score (CES) that is used in connection to customer support and shows how easy it is to get a problem solved.

Due to the immense use of the NPS-score in the SaaS business, this score is worth some more explanation. The scoring model is a result of a big research attempt to find out what metric that could be used for best gauging customer loyalty with the earlier mentioned Frederick Reichheld (2003) as main researcher. The research resulted in a scoring model built around one question: “How likely are you to recommend the product or service to a friend or colleague?”, since this question was found to have the strongest correlation with repeat purchases and referrals (Reichheld, 2003). The score is measured on a scale from zero to ten, were ten equals “Extremely likely”, five equals neutral and zero equals “Not at all likely”. After studying customer ratings and repurchase behavior, the researchers found that the responses formed a pattern with three distinguishable clusters: “Promoters” giving a nine or a ten, “Passively satisfied” scoring a seven or eight, and “Detractors” scoring from zero to six. The final score is calculated by subtracting the percentage of “Detractors” from the percentage of “Promoters”, in that sense excluding the lukewarm middle and avoiding the grade inflation that according to Reichheld (2003) often occurs in satisfaction measures and is the result of labeling any customer that score above the neutral level as satisfied. The strength with the NPS-score is not only that it is a fairly harsh scoring method, it is also assumed to separate the true loyal customers from the ones who repurchase because of reasons such as inertia. Reichheld (2003) argues in the following way: “loyal customers talk up a company to their friends, family, and colleagues. In fact, such a recommendation is one of the best indicators of loyalty because of the customer’s sacrifice, if you will, in making the recommendation.”

To summarize, the satisfaction-loyalty-profit chain is a concept that is central in the field of Customer Relationship Management even though the different parts of the chain has been subject to great scrutiny from both academia and practitioners. However, the general notion is that improved customer satisfaction and loyalty will eventually lead to greater company performance such as increased profits. Scoring methods such as CSAT and NPS has therefore
been used to gauge how customers are doing and is also used as input for additional marketing initiatives to increase customer retention and profitability. At this point, it should be clear that firms could yield several benefits by focusing on customer retention rather than investing heavily in acquiring new customers. For that reason, it is equally important to minimize customer defection. This is a concept that is well studied and described in the CRM-literature, and a concept that has a strong linkage to this study. The means for predicting customer defection will be more closely described in the following section.

2.4 Customer defection and churn prediction models

Customer defection, or customer churn, is a phenomenon that most certainly affects every business. As earlier mentioned, churn is defined as “the tendency for customers to defect or cease business with a company” (Kamakura et al., 2005), and is an event that can turn out to be rather grave if the company’s churn rate is higher than the acquisition rate. Many companies have for that reason emphasized the need to find out what warning signals that exist, to identify what customers are likely to defect in the relatively near future. This information can thereafter be used to send out incentives to the right customers, in order to make them stay. A solution to this problem is the use of churn prediction models, which aim is to determine the drivers of customer churn and, when applied on an existing customer base, point out what customers that are in the risk zone of leaving the company. Because the decision to churn is a binary decision, logistic regression is a popular method for developing such models. Other methods used have their roots in data mining and machine learning, such as neural networks and random forests (Kumar & Petersen, 2012, p.152). There exist plenty of research on what modeling techniques that leads to the best predictive models, however the results are often industry or company specific and depends very much on the data being studied.

The general process when developing a churn prediction model with the help of data mining starts with defining what business objectives the model should aim at reaching. The subsequent steps include extracting raw data, identifying relevant variables, gaining customer insight and finally acting upon the results (Kumar & Reinartz, 2012, p.144). This process is shown in figure D.

![Figure D](image-url)

Figure D: An overview of the data-mining process. (Own, after Kumar & Reinartz, 2012)

Although the creation of a churn prediction model goes beyond the scope of this study, parts of the process can be looked at and imitated in order to identify customer health parameters.
and eventually develop a customer health score. This study will therefore borrow the general procedure explained above, with some modifications. The study will for example not include the step of testing and training different predictive models, which is the general process in most churn prediction studies. However, logistic regression will be used as a method to find out what parameters that best describe the likelihood of defection, parameters that might be weighted heavier in a final customer health score. The next section will look at what parameters previous research has found as important when developing predictive models, information that can be used as a starting point for this study.

2.5 Predictive parameters used in previous studies

One of the early steps in creating a customer churn model is to identify what data that can be extracted as well as defining what parameters that might have the ability to indicate a potential churn candidate. Ng and Liu (2000) suggest that usage data should be used for recognizing churn in the Internet service provider and telecommunications industry. A service industry that works with subscription-based contracts and in that aspect is similar to Funnel’s business model. Another study by Szucs and Kiss (2013) used usage data such as time spent in the tool and number of log-ins as parameters when developing a churn prediction model for a software provider in the mobile applications industry. Similarly, to refer back to the SPC-literature, research by Bolton and Lemon (1999) and Ram and Jung (1991) shows that satisfied customers are more likely to have a higher service usage level than unsatisfied customers.

Regarding predicting customer patterns such as future purchases, there have been claims that historical purchasing data is a good indicator for such behaviors (Verhoef & Donkers, 2001). A large amount of studies suggests that Recency, Frequency and Monetary (RFM)-variables are good parameters to look at when predicting customer behavior (Hsieh, 2004; Liu and Shih, 2004a; Jonker et al., 2004; Verhoef & Donkers, 2001). Liu and Shih (2004a) describes the respective variables in the following way:

**Recency**: variables indicating time since last purchase or use of a service. According to the authors a lower value suggests a higher probability that the customer will make another purchase.

**Frequency**: variables related to how often the service is being used. According to the authors a higher value indicates higher loyalty.

**Monetary**: variables that show how much money a customer has spent over a certain amount of time. The higher the value the more important the customer should be to retain for the business.

These three variables have been heavily used for all sorts of predictive analyses, such as providing product recommendations in e-commerce based on customer lifetime value (CLV) (Liu & Shih, 2004a), predicting future partial customer defection in non-contractual fast-moving consumer goods industry (Buckinx & Van den Poel, 2005) and predicting customer churn in the online gambling industry (Coussement & De Bock, 2013), just to name a few. Furthermore, Wu and Chen (2000) has noted that the more recent a customer has purchased a
product the higher the probability that the customer is active, and the frequency of purchased products can be a measurement for the likelihood of the customer churning in the future (Reinartz & Kumar, 2000). In other words, these three variables have been confirmed to be of utmost importance when predicting customer churn (Buckinx & Van den Poel, 2005; Coussement & De Bock, 2013). It should be said that research on churn prediction models based on data mining techniques has been mainly applied to business to consumer (B2C) contexts. One reason is that the use of big data has yet not made itself a big name in the B2B-setting, as is the case in the B2C (Wiersema, 2013). However there have been some attempts to develop predictive models also for B2B companies such as the work by Tamaddoni Jahromi, Stakhovych & Ewing (2014) in which RFM-variables were used as predictive variables.

To conclude, as indicated by the presented literature, parameters such as usage data and RFM-variables could be interesting to look at in order to figure out what distinct parameters that might impact customer retention and churn for Funnel’s customers. In addition, customers’ level of satisfaction and loyalty might also be revealed in their usage behavior in such a way that a satisfied customer uses the tool more than someone who is less satisfied. Lastly, theory linked to the SPC-concept can work as background when defining customer health for Funnel’s business, and customers’ level of satisfaction, loyalty and retention can preferably be measured by sending out questions similar to the ones in the NPS- and CSAT-surveys.
3 Method and execution

This chapter will go over the methods that have been used in this study and describe how the research has been conducted. The chapter will begin with a description of the study design and strategy, and thereafter describe each method that have been used for collecting data, as well as for analyzing. The chapter will then end with a critical evaluation of the methods chosen, as well as give a short description of the criteria in research and how these criteria have been taken into account in this study.

3.1 Study design and strategy

The purpose of this study was to extend the existing knowledge of customer health as well as to show how one might tackle the problem of identifying relevant parameters to measure this concept. This required an intensive examination of the company that was the object of study and therefore a case study design was seen as suitable. In addition, this study’s process has mainly had an inductive approach. The difference between a deductive and inductive approach is according to Bryman and Bell (2011, p.13) the different relationship between theory and empirical evidence. While a deductive approach is often linked to quantitative research and characterized by theory being the starting point of research on what one or several hypotheses are based, an inductive stance is more commonly used in qualitative research and views theory as the outcome of research. In other words, inductive research often includes drawing generalizable conclusions out of the study’s observations (Bryman & Bell, 2011, p.13). It is important to mention that although deductive and inductive approaches seem like each other’s antipoles, the distinction between them is far from clear-cut. Deductive strategies often include elements of inductive aspects, and the opposite goes for inductive strategies (Bryman & Bell, 2011, p.14).

With that in mind, this study has mostly applied an inductive approach with several qualitative methods being used, however the outcome has been little more than empirical generalizations and the study has also included deductive aspects such as using theory as background to the study’s execution. The process itself has been iterative, meaning that there has been a continuous weaving back and forth between data collection and theoretical reflection (Bryman & Bell, 2011, p.13). Previous research and theory has for example worked as an introduction to the case at hand and suggested what parameters that might be of interest to study. Subsequent methods have thereafter complemented this initial information and narrowed down the research to only focus on the parameters that are suitable to examine for the company studied. In such a way, each method has ended with a theoretical reflection, which have worked as a starting point for the collection of more data by the use of additional methods. That being said, some of the methods that will be described in the next sections have been conducted parallelly, while some have overlapped each other. In addition, some of the parameters have been redefined several times to improve their validity, which means that the methods have not been conducted totally back to back. This approach has given the researcher flexibility and allowed the methods to be adjusted slightly during the study.
3.2 Choice of methods
The methods that were chosen for this study were of both qualitative and quantitative character and consisted of several interviews, participant observations, a web survey and a data mining execution. Since the study centered on understanding customers and their usage patterns, the need for more in-depth customer insight was apparent from the start. Therefore, the study began by taking a more interpretive position and employing more qualitative methods such as qualitative interviews and ethnography. In interpretivism, the focus lies on understanding the social world by studying and observing how people make sense of, and interpret, that world (Bryman & Bell, 2011, p.386). By conducting qualitative interviews and spend most of the time at Funnel a first understanding of the tool as well as product usage could be obtained. This was also the aim of using the two different research methods early on in the process. However, both of the methods were continuously used in latter parts of the study to get a better and more complete picture of the case at hand.

As mentioned, quantitative methods were also used in the study, and followed the initially obtained qualitative data. According to Bryman and Bell (2011, p.163) the preoccupations of quantitative research can be summarized as dealing with measurement, causality, generalization, and replication. Because of this focus on measurements and causality, and quantitative methods’ ability to produce data that is easily measured and compared, the use of quantitative methods seemed suitable for this study. The specific aim with employing such methods was therefore to be able to measure customers’ product usage, parameters linked to their accounts, and survey results, measurements that afterwards could be used in quantitative analyses to try to find correlations between parameters. Quantitative methods were in other words necessary to include in the study in order to answer research question four.

The sampling frame that was used in both the qualitative and quantitative methods mainly consisted of Funnel’s customer base, including both current and previous customers. Each method with its respective methodology and analysis will be described in greater detail in the following sections as well as the type of sample chosen for each method.

3.3 Interviews
Several interviews were conducted in this study with the aim of reaching customer insight as well as understanding the daily customer work at Funnel. The decision to choose interview methods to get such insight was supported by Alvehus (2013) who argues that interviews are one of the most effective methods to gather data about how people think and feel. Furthermore, qualitative interviews were chosen as the most suitable interview method since it puts a bigger interest in the interviewee’s point of view, while quantitative interviews more often reflect the interviewer’s (Bryman & Bell, 2011, p.466). Qualitative interviews also tend to be less structured, which were seen as an advantage in many ways. The possibility for the interviewee to elaborate on topics was seen as one. Another one was the interviewee’s possibility to go off at topics linked to the one discussed, since this gives further insight into what the respondent views as important and relevant (Bryman & Bell, 2011, pp.466-467). From the interviewer’s point of view a qualitative approach allowed to follow up on replies that were seen as especially interesting for the study and change the order in which the
questions were delivered. Qualitative interviews are also good at receiving rich and detailed answers that can give a deeper understanding of the respondent’s world (Bryman & Bell, 2011, p.467), which were seen as a big advantage when aiming for a deeper customer insight. With this as background, semi-structured interviews were seen as a preferable method since such an approach follows an interview guide with pre-defined questions or topics. This would make it easier to afterwards compare the answers and distinguish characteristics and similarities in respondents’ answers (Bryman & Bell, 2011, p.467). The interviews were conducted with both employees at Funnel as well as their customers. For the sake of simplicity, each interview round will be described separately in the following sections.

3.3.1 Semi-structured interviews with Funnel’s employees

In the beginning of the study, semi-structured face-to-face (F2F) interviews were conducted with four pre-chosen employees at Funnel. The employees were chosen with respect to their field of responsibility and therefore included personnel from both Customer Success and Product Development. The aim of these interviews was to get a first understanding of how the company’s current customer work looked like as well as getting input regarding the concept customer health; how it should be defined and measured. Therefore, the interview template was divided into sections including questions about general customer work, customer health, creation of a customer health score, and product usage. The interview template can be found in appendix A.

3.3.2 Semi-structured interviews with Funnel’s customers

The aim of the customer interviews was to understand the customers’ world: how they use the tool as well as their thoughts and experiences of both tool and customer service. Therefore, an interview template was designed covering areas such as product usage, product value, and company support, see appendix B. The interviews were either conducted F2F or over telephone with the help of the instant messaging app Skype.

Regarding the sampling methodology, a non-probability sample was chosen. In other words, participants were picked with respect to how likely they would be to accept an interview invitation. Churned customers were for example not interviewed, since it was assumed that these customers would be hard to get in contact with and would view the opportunity as hardly beneficial. Instead, the interviews were solely conducted with active customers, and were conducted in two different rounds. The difference was that the first round purely consisted of customers seen as promoters by Funnel’s staff, an aspect that was given less importance in the second round. The first sample can therefore be defined as a pure convenience sample that was picked in order to quickly get customer input to the research. The second sample was picked with consideration to additional factors such as customer segment, product segment, and usage pattern. By trying to pick customers that differentiated from each other and to a higher degree resembled the diversity of the actual customer base, one can argue that a kind of quota sample was being used. However, the sample that has been studied has been far from random and might include sampling errors, aspects that will be examined in greater detail in the section Critical evaluation of method.
3.4 Ethnography

It was early on recognized that this study would reap several advantages by having the researcher based at Funnel when working on the thesis. Such advantages were for example the possibility to ask employees for help on anything from information on specific customers, to questions about how to extract data from the company’s database. It was also necessary in order to understand how the company’s product works as well as getting a better understanding of how customers use it. Therefore, the researcher immersed himself in the company and in that way conducted something that would be best described as a participant observation or ethnography. Although the objective was not to study Funnel’s employees per se, the decision to spend most of the time at the company allowed the researcher to observe and learn about routines, ask continuous questions, and get updated information about what was going on in the company at the moment. Without that possibility, the study would have been much harder to conduct and handle. One should remember though, that while an ethnography allows the researcher to get information that otherwise would have been hard to collect, there is always a risk of “going native” and thereby losing objectivity.

3.5 Customer web survey

As mentioned, quantitative analyses were seen as necessary in order to examine the big amount of product usage data that Funnel collects daily. However, it would have been hard to get a grasp of the customers’ own perceptions of the product without including quantitative responses from the customers themselves. For this reason, a web survey was chosen as a suitable method. The aim was to get quantitative data that reflected the level of satisfaction and loyalty among the customers as well as their view on the customer support. However, some qualitative questions were also added in order to better understand some of the customers’ ratings. Apart from the fact that the survey data could be included in the planned quantitative data analysis, the method had several other advantages. Firstly, a survey is cheap to administer since no time and money have to be spent on travels or telephone calls. Secondly, it is quick to administer since questionnaires can be sent out in batches, reaching a bigger sample in less amount of time compared to telephone or F2F interviews. And thirdly, it is convenient for the respondent, since he or she can answer whenever they have the time. However, it is worth to mention that this convenience comes with a risk of delayed, forgot or unanswered responses, something that might be partly tackled by sending out reminders (Bryman & Bell, 2011, pp.232-233). The survey was designed as an online self-completion questionnaire with the help of Google Forms and mainly consisted of questions that were answered by giving a measure on predefined scales. It was sent out to customers by mail together with a short description of the aim and content of the survey. The web survey template can be found in appendix C.

The sample chosen for the customer survey was created in Funnel’s software tool for customer communication: Intercom. By using the filtering function the mail could be specified to only be sent to customers that 1) signed up more than 45 days ago, 2) used the tool less than 30 days ago, and 3) had used the tool more than 10 times. The reason for creating such a filter was the desire to reach customers that had been paying customers for at least one month, used the tool within the last month (in that sense targeting relatively active
users that used the tool on a regular basis) and used the tool more than just a few times so that they would have had the time to understand the tool and its functions, as well as had the chance of contacting support.

3.6 Data mining execution
Since Funnel is a SaaS company that provides its software tool as a cloud service there existed great opportunities for investigating the data that Funnel’s customers create through their daily usage. A big part of the product usage is monitored and stored which means that there existed current as well as logged historical data, for both active and churned users. Furthermore, the stored data is on user level, which means that not only was there usage data for each company, but for every user at each company that used the tool.

The aim of studying this quantitative data was to find what parameters that best indicated whether customers were of good health or not. Both researcher and company were of the opinion that the database included information that could give a better understanding of how the customers use the tool as well as how different usage patterns could be interpreted as different customer behaviors. To be able to assemble this information a data mining execution was chosen as a suitable method, which means that quantitative data was extracted from the company’s database. The aim of the method was to extract data about different parameters so that these later on could be analyzed in the statistical analysis software IBM SPSS Statistics as well as compared to information from both web survey and interviews. The expectation was that the quantitative data from the data mining execution would show the “real” customer behavior, complementing the stated behaviors from the interviews and the customer ratings from the web survey.

The methodology itself consisted of writing queries into Funnel’s database using the domain-specific language SQL, which stands for Structured Query Language. What parameters to study were based on information coming from the employee and customer interviews. The process of extracting information about the parameters then began with defining each parameter. Thereafter, the database, consisting of data stored in different tables, was examined in order to understand what tables that needed to be combined to build up the parameter and also how to relate this data to each company. After that a SQL-query was written that would call each specified table in the database, collect the relevant data and assemble all information in one new table showing the value of the parameter for each company. This data sheet was thereafter downloaded as a .csv-file, ready for further analysis.

As described, the sample for the data mining execution was both active and churned users. Regarding historical data, a limit was set to the beginning of year 2016, partly because the amount of data back to that date was seen as enough for the study, partly because the tool had undergone so much change in the last year so that looking further back in history would have given a misleading picture of customers’ product usage. Further limitations were taken concerning what kind of customers to look at. Trialing customers were for example excluded from the sample with the explanation that these shouldn’t be viewed as real customers, partly because they might deviate in their product usage from more experienced users, partly
because they might use the tool with less commitment due to testing the tool for free. Another limitation was the exclusion of Funnel’s own staff, since they aren’t part of the customer base and therefore shouldn’t be studied when trying to define customer health parameters. Other than that, the sample within the specified time limit was seen as the whole population, resulting in a probability sample with some limitations.

3.7 Analysis of qualitative content

The material coming out of the interviews as well as the qualitative question in the web survey were all object for qualitative analysis. Beginning with the interviews, each interview was recorded and later on transcribed in order to make the analysis easier. Yin’s five-phase cycle for analyzing qualitative data (2011) were thereafter used in order to structure the material and open it up for further analysis. The cycle, as well as how it was applied to the study’s material, is described in greater detail below:

1. **Compiling:** The five-phase cycle begins with compiling and sorting data, so that the assembled data is structured in some way. This was done in such a way that each interview transcript was put in chronological order and placed in a shared data folder.

2. **Disassembling:** In the second phase the compiled data is broken down into smaller fragments. For this step Alvehus’ (2013) method of color-coding was used since it appeared to present a fairly time-efficient and easy methodology. In practice, this meant that each respondent’s answer on each interview question was simplified and given a color-code that referred to that specific interviewee. Each question’s simplified answers were thereafter grouped together, to allow for further comparison.

3. **Reassembling:** After the disassembling phase the fragments are reorganized into clusters to make it easier to distinguish the different themes. This phase can be repeated with the previous one in order to identify more accurate themes. In this study, similar characteristics were identified among the color-coded answers and clusters were formed according to these similar themes.

4. **Interpreting:** In the fourth phase quotes or narratives are presenting the themes that have been identified in the previous section. It is also possible to present the information in the form of graphs and tables. These narratives form the basis of the empirical findings. In this report, the distinguished themes are presented in text as well as with the use of quotes.

5. **Concluding:** In the last phase of the cycle the final conclusions from the study are drawn, relating to all the previous phases.

A similar procedure as the abovementioned was followed for the qualitative survey question, in which the answers were first grouped into groups according to how the respondents had rated their likelihood to stay as a customer. Common themes were thereafter identified and grouped into clusters by the help of color-coding, ready to be further interpreted together with the rest of the study’s results.
### 3.8 Analysis of quantitative content

This section will describe the procedure behind the analysis of data coming from the data mining execution. As mentioned, a .csv-file including information about each company and their parameter values were extracted from the database, ready for further processing. The statistical analysis software IBM SPSS Statistics was chosen as a suitable program for conducting the statistical analyses required to study the relationships between parameters. The procedure linked to this will be described in the next section.

To begin with, the .csv-file was opened up in SPSS, in order to clean the data. This means that variables were given labels and were defined according to their variable type, missing values were identified and defined, string variables were recoded into numerical values, and variable values were given explanatory labels. Apart from this common cleaning procedure, the data was further altered in such a way that outliers three standard deviations away from mean-values were detected and removed, some of the scale variables from the survey were computed into shorter scales, see Appendix D, and filter-variables were created in order to filter out cases that were not going to be part of the analysis. After this procedure had been done, the data was ready for analysis.

The analyses that were chosen for this study were crosstabulation, one-way test of variance (so called ANOVA), linear regression, and binary logistic regression. Explained shortly, a crosstabulation test, or a contingency table, allows the researcher to examine two variables simultaneously, so that relationships between the two variables can be analyzed. The result is illustrated in a matrix format where the real count in each cell can be compared to the expected count, in that way indicating if there exist any salient interrelations between the variables (Bryman & Bell, 2011, p. 347). The ANOVA test is a statistical method used to analyze variance, from which inferences about mean values can be drawn. The results of the test show whether the mean values for several groups are equal or not, and can therefore be used when analyzing differences between two or more groups (Lærd Statistics, 2013). Lastly, the regression analysis is a popular predictive modeling technique. The aim of the analysis is to examine if, and how strongly, one or several independent variables affect a dependent variable, in other words to examine the correlation between variables. In linear regression analysis, the relationship between the independent and dependent variable is approximated by a linear line, the regression line, which shows the best explanation of the correlation. The relationship is further explained by a regression formula, indicating the start value when the independent variable equals zero, the slope of the line, as well as the error term. In binary logistic regression analysis, the dependent variable is binary in nature, for example: 0 or 1, true or false, yes or no. Since the distribution is binominal, the link function that is used is the logit function. The relationship between the dependent and independent variables is expressed by a regression formula, as in linear regression (Ray, 2015).

By using these different statistical methods, the relationships between the web survey variables, as well as parameters from the data mining execution, could be examined. Before each test, variables were chosen and then defined as either dependent or independent. The
output of each test was thereafter examined together with values on statistical significance. The results from each test can be found under Empirical results.

3.9 Critical evaluation of methods

3.9.1 Interviews
Although each method in this study has been conducted with a certain aim in mind and checked several times for errors and ambiguities, there is a risk that the chosen methods have included some kind of deficiencies that have affected the results in an unexpected way. Starting with the customer interviews, each interview was only conducted with one person at each company using the tool. When that person is the only person using the product that is not a problem, but when more than one person at a company is using the tool there is a risk that the person being interviewed has a different opinion than the others. Resulting in a result that is not totally representative for that company.

Other errors linked to the interview process are misunderstanding of questions, as well as bias linked to both interviewer and interviewee. An example of such a bias is that the interviewees might have answered differently than they would normally do because they have perceived the interviewer as representing the company, therefore responding in a more positive way or leaving out pieces of information that would have been relevant for the study.

3.9.2 Customer web survey
Moving on to the web survey, there exist four different kinds of errors in survey research according to Bryman and Bell (2011, p.196). Namely sampling errors, sampling-related errors, data collection errors, and data processing errors.

To begin with, it is very unlikely that the respondents to the web survey could be seen as a truly representative sample of the whole customer base, which increases the risk of sampling errors. Far from everyone that received the survey answered it, and there is a potential risk that the ones that answered it had some characteristics that made them more likely to answer. For example, there is a risk that respondents have belonged to two extremes, the ones who love the tool and therefore want to express their satisfaction, and the ones who dislike it and wants to make their voice heard, in other words leaving out the big mass in the middle that stay lukewarm to the product.

Regarding sampling-related errors there is a risk that the sampling frame for the web survey has been inaccurate. As already mentioned, the sampling frame was constructed using a filter in Funnel’s customer communication software, and there is a risk that the use of that filter resulted in a non-representative sample for the survey. To this category also belong non-responses, which have already been discussed in the previous section. Far from everyone that received the survey answered it, which of course have affected the result of the survey.

The next error to discuss is data collection errors. To this category belongs problems linked to poor wording and ambiguity of survey questions, a risk that is always present in survey
research, especially when the survey is constructed as a self-completion questionnaire and the respondent have little opportunities to ask for clarifications of specific questions. There is for example a risk that respondents have interpreted the scales or questions differently, and therefore answered them in a faulty way.

Lastly there are data processing errors, in other words errors related to the coding of survey answers. However, the survey mainly included closed questions, which made the coding easy and reduced the risk of such errors. The few qualitative questions were chosen not to be included in the quantitative analysis and therefore didn’t undergo any coding, eliminating eventual data-processing errors linked to these answers.

3.9.3 Data mining execution
There are mainly two errors linked to the data mining execution that might have occurred in this study. Firstly, there is a risk that the queries that were written were faulty, resulting in parameters that aren’t showing the right thing. Secondly, there is a risk that the matching of data in the database was performed in a wrong way, resulting in data being linked to the wrong user or customer. However, these two errors tried to be avoided by test-running the queries multiple times and double-check for errors, as well as single out certain users to validate that the right information belonged to the right user.

3.9.4 Analysis of quantitative content
The analysis of quantitative content has also been prone to errors. Firstly, there is a risk that data have been handled in the wrong way during the cleaning-phase for example by not labeling missing values correctly or by recoding variables in the wrong way. Secondly, there is a risk that the data have been altered in a wrongly manner, for example when removing outliers. Such inference by the researcher should be performed with great caution and only if necessary, since such conduct is totally dependent on what the researcher view as good and poor data points. Thirdly, there is a risk that the data chosen for the analyses have been of the wrong variable type or wrongly distributed, resulting in results that aren’t trustworthy. Lastly, the analysis of results could also have been subject to errors, for example noting statistical significance when there is none.

3.10 Criteria in research
Reliability, replication, and validity are the three most important criteria for evaluating the quality of a research study. Reliability refers to the question of whether the results of a study are repeatable or not, and if the measures being used are consistent (Bryman & Bell, 2011, p.41). In other words, would the same results be obtained if the study was repeated, or have they been generated by chance? Replication, closely linked to reliability, is concerned with the replicability of a study. In other words, whether the study itself is repeatable. For a study to be replicable it is important that all procedures are accurately documented (Bryman, 2002, p.49). Lastly, validity is an evaluation of whether the conclusions that are generated from research are coherent or not (Bryman, 2002, p.50). The concept of validity can be further divided into:

- Measurement validity: Are we measuring what we think we are measuring?
• Internal validity: Do the proposed causal relationships hold water?
• External validity: Can the results be generalized beyond the research context?
• Ecological validity: Are the findings applicable to people’s everyday life?

All these mentioned criteria have been taken into account during this study. To ensure the reliability of the study, the theory, sample and chosen methods have been documented and motivated in this report. However, the results are dependent on the sample that have been chosen for this study, and might therefore differ slightly if the study would be repeated. Regarding the study’s replicability, the connection between theory and empirical results has been documented as well as the implementation to ensure that the study can be repeated. However, it is important to note that the qualitative parts of this study might be harder to repeat, since the implementation and analysis linked to these qualitative methods are prone to subjectivity (Bryman & Bell, 2011, p.408). Finally, the study’s validity has been taken into account by describing the conclusions in detail with connection to both results and theory. The internal validity might be most important since the study set out to study relationships between variables. To ensure internal validity, caution has been taken when drawing conclusions about casual relationships in the material.
4 Ethics

Ethical issues related to research are almost impossible to completely eliminate, and that is why a discussion of ethics always should follow a study like this. According to Bryman & Bell (2011, p.143) the boundary between ethical or unethical practices are seldom clear-cut, and for that reason several codes of ethics have been created over the years. These documents, created by both professional associations and the social science field, have been produced in order to give researchers guidance in dealing with ethical questions. For this study, the Code of Conduct and Guidelines by MRS (2014) has been used as a reference guide. The reason for choosing this guideline was its comprehensiveness of dealing with both quantitative and qualitative research as well as its suitability for this study in terms of what methods that have been used.

4.1 Ethical principles

To begin with, ethical transgressions can be divided into four main areas, as done by Diener and Crandall (1978), namely: the risk of harm to participants, lack of informed consent, invasion of privacy, and deception. These four areas will be discussed in the following sections as well as issues linked to data management, since the handling of data has been a central part of this study.

Starting with harm to participants, according to the Code of Conduct and Guidelines by MRS (2014, p.3) “Researchers shall ensure that participants are not harmed or adversely affected by their professional activities.” This case study hasn’t included any practices that could have led to physical harm to the participants, and the only practice that could have led to psychological harm would have been the sharing of confidential information from interviews. However, no such information has been shared in this report and as a general rule of thumb all participants were given anonymity (MRS, 2014, p.15).

Regarding the ethical principle of informed consent, Bryman and Bell (2011, p.133) states that potential research participants must be given enough information needed to make an informed decision about whether they want to participate in a study or not. Likewise, such need for informed consent is also outlined in the MRS Code of Conduct and Guidelines (2014, p.13). These guidelines have been followed in this study by clearly informing the prospective participants about the mission of the study when asking for their participation. In addition, all participants were notified of any recording equipment that was used during the interviews (MRS, 2014, p.17).

Regarding the privacy of individuals, researchers should take careful steps when asking for sensitive information and make sure that the respondents don’t feel forced to share information that he or she might view as of personal character (Bryman and Bell, 2011, p.136). In order to avoid such invasion of privacy the interview questions in this study were carefully chosen so that they shouldn’t intrude on any respondents’ privacy. To further protect individuals’ privacy, personal information has been kept confidential and anonymous in this report. The collection of user data from Funnel’s database is also a case linked to individuals’
privacy. However, no information of personal character was assembled. Additionally, Funnel’s customers had given their consent to their data being shared within the company and that it might be used for developing the product, a process that could also be argued to be the aim of this study.

Deception, according to Bryman and Bell (2011, p.136), “occurs when researchers represent their research as something other than what it is”. Therefore, in order to avoid deception being part of the research, the mission of this study was clearly described for all participants.

Regarding the storing and handling of data, all data assembled were stored digitally. Although little to none of the data that were assembled could be seen as harmful, precautions were taken to ensure that data were stored as safe as possible. Data regarding customers’ product usage as well as survey answers that could be linked to specific individuals were stored in the company’s cloud, and were in that sense kept within the company’s boundaries. This follows the MRS Code of Conduct and Guidelines (2014, p.20) that states that: “Members must take reasonable steps to ensure that all hard copy and electronic files containing personal data are held, transferred and processed securely in accordance with the relevant data retention policies and/or contractual obligations.” However, information such as recordings and transcriptions of interviews were also stored on the researcher’s personal computer, a procedure that can be questioned from a security perspective.

4.2 Further ethical considerations
Although most ethical principles related to the execution of this study has been mentioned, there exist ethical considerations that not directly falls within the conduct of this study worth mentioning. These ethical considerations are all linked to customers’ financial data stored by the company. Because even though financial data such as advertising spend and budgets have not been included in this study, the researcher had access to this information when working on behalf of the company. There are at least two practices that could be seen as unethical related to this. Firstly, to share confidential information with people outside the organization such as customers’ advertising spend or budgets would be seen as highly unethical. Secondly, to use this information for insider trading activities in case companies are listed on the stock exchange could be seen as illegal, and highly unethical. Both these practices were avoided and therefore absent during this study.

4.3 Wider ethical implications
Talking about ethical implications in a wider sense, this study can be seen as incorporating practices such as monitoring users and storing their data. Starting with the practice of monitoring customers, this is a commonly used practice in the SaaS industry. The reasons for monitoring customers are plenty, but product development is one of importance for Funnel. By continually monitoring customers the development team can for example see how new features are welcomed and in that sense, get continuous customer feedback for further product development. However, the practice is not only of help for the company. In the case of service failure or when problems arise with the product usage, the customers have the option to call in to support. In such a scenario, it is of importance to know whether the error occur because of
the customer’s wrongdoing or if there is something wrong with the service. Without having the ability to quickly gather data and analyze it, troubleshooting becomes difficult. Therefore, the monitoring also benefits the customers. However, the knowledge of being monitored might fill some people with unease, and that is a question of ethics. It might for example be argued that this practice infringes on the individual’s privacy, even though the data monitored doesn’t include any sensitive personal data. It might also be argued that customers should have the possibility to choose whether or not to be monitored, but since this would make it impossible to help a customer in a situation of failure, and would complicate the support and set-up process, it is very unlikely that such an option would be possible to offer, or even desirable for the customer. In the end, the practice of monitoring is debatable from an ethical standpoint, even though in this case most customers probably would value the benefits higher than the sacrifice of privacy.

Linked to the monitoring of customers is the practice of storing and transferring customer data including financial information about each company, such as advertising costs and budgets. Such data is important to store for Funnel in order to offer backups if the service goes down or if there are any problems with the connections to external sources. Furthermore, the extracting and transferring of financial data from each advertising platform as well as data from Google Analytics is what builds up the product, and without that functionality the tool wouldn’t work or deliver any value to its customers. However, ethical implications might arise if this information isn’t stored or transferred in a secure way, threatening its confidentiality and possibly the customer’s business. To avoid such consequences great weight lies on secure storage as well as transfer of data. Further handling and disclosure of confidential data within Funnel is specified in a Non-Disclosure Agreement, signed by both Funnel and the customer. Unethical actions such as disclosing customers’ confidential information to people outside the organization is therefore regulated by this document and seen as a breach of the agreement with injunctive reliefs as consequences.
5 Empirical results

This chapter will present the empirical results coming out of the earlier mentioned research methods and will include enough information to answer all of the remaining research questions. Since research question one already has been brought up in the Theory and Literature chapter, it won’t be included in this chapter. Research question two and three, however, will be possible to answer by Results from employee interviews and Results from customer interviews, while research question four is rather linked to the Results from customer web survey and Results from data mining execution. For the ease of reading each research question will be written out in the beginning of the relevant section.

5.1 Results from employee interviews

Research question 2:

What is seen as good and poor customer health at Funnel and how can this be identified in customers’ product usage?

Ten semi-structured interviews were conducted out of which four were with Funnel’s employees and six with customers. From the employee interviews, it turned out that the company nowadays lack efficient ways of quickly getting an overview of how their customers are doing and recognize customers that are in the risk zone of churning. The amount of time spent on different customers seemed to vary a lot, with the general feeling that the “loudest” customers get the most attention. The ability to allocate resources and time where needed was therefore mentioned as one of the greatest benefits with developing a customer health score. When being asked about how customer health should be defined the following definitions were mentioned by several of the employees. Namely customer health is:

- A measure of the quality of the relationship between company and customer
- A measure of the perceived value that the customer gets out of using the product

In addition, the interviewees mentioned that a good customer health should equal a customer that is an advocate for the firm, having low risk for churn and being responsive for upsells. A poor customer health, on the other hand, should indicate an impending risk for churn, or even worse, a customer that is the opposite to an advocate:

A risk with poor customer health is that [the customer] defiles our brand in different ways, like writing bad reviews or in other ways spread a negative message about us. – Funnel employee, 2017-03-09

Regarding what parameters that possibly could be the most important to look at, the following were mentioned as good candidates: product usage (such as average of log-ins), number of active users, support tickets, number of data sources and the NPS-score. However, one mentioned risk with looking at usage parameters such as activity in the tool, were the ambiguity in interpreting the parameter. A high product usage could for example be interpreted as the customer getting a lot of value out of the tool and for that reason using it regularly, while it could also mean that the customer is struggling and spends a lot of time in the tool because of confusion. The interviewees also stressed the importance of knowing what
type of customer you are studying and that the usage pattern could differ between the different product segments.

Furthermore, several of the employees mentioned that customers that early on establish a frequent usage pattern after their trial setup were less likely to churn. This had been shown in a study prior to this one, which had examined Funnel’s customer base and investigated the differences between a successful and a non-successful customer. This further implied that usage parameters, such as usage frequency, could be interesting to look at. The heterogeneity among customers was further stressed by an interviewee who meant that the product use very much depend on the individuals using it. According to him an individual working as an e-commerce manager probably used the tool differently than a person working as a digital marketer. Finally, when asking if there existed any “sticky features” that possibly could indicate that the customers are relying heavily on the tool and in that sense, are more inclined to stay, the sharing of reports were mentioned as one possible. As one of the interviewees said: “if you share something with someone, then you are often satisfied with what you have created.”

5.2 Results from customer interviews

Research question 3:  
What value is the product providing its users and how is this related to customer health?

Out of the six customer interviews, three were with e-commerce companies, two were with agencies, and one with a SaaS company. Five of the interviewed persons used the tool almost daily, while the last interviewee used it only twice a week. The reasons for using the tool in the first place could be grouped into three main categories: forecasting, reporting and situation assessment. Regarding the situation assessment one of the agencies described the tool as their “single source of truth”, a description better explained by the following quote from the same customer:

Is the money spent better here, or is it better spent over there? It’s the only place where we answer that question. – Agency customer, 2017-03-14

This overview functionality was also mentioned by the SaaS customer who said: “It’s unfortunately one of those boring products that sit in the background, but at the same time you rely on it, so when you look at it you know it’s done.” Apart from the abovementioned reasons, every respondent also mentioned saving time as a benefit with using the tool. Regarding the customer support, all of the interviewees mentioned they were satisfied with the support given and found the ease of contacting good. One interviewee even mentioned that they often didn’t have to contact support:

Most of the times the support-team tell us about a problem before we have experienced it ourselves. – SaaS customer, 2017-03-24
Just like during the employee interviews, the customers were questioned about what parameters they thought would best indicate good and poor customer health. Parameters mentioned by several of the respondents were: activity parameters (such as time spent in the tool, number of log-ins and checking in on ad spends), the amount of sent reports, and individual parameters decided during set-up. Other parameters mentioned were amount of support tickets, changes in account size and sharing features. Regarding what in the tool that gave the customers value saving time was mentioned as a big one. Another person expressed the surprise when a new functionality gave them additional value that they hadn’t expected and one year later gave them the means to solve problems they otherwise wouldn’t have been able to solve. Another interviewee saw the potential of getting more value out of the product, if they could get more people in their business to use it. However, the same person also mentioned the importance of improving sharing features in the tool in order to get more people in each business to use it. The following quote describes his reasoning well:

I think the whole sharing feature could be much better, […] basically every SaaS product is always very keen on you inviting other team members, cause that is always what helps you stick. – SaaS customer, 2017-03-24

However, it is worth mentioning that all of the interviewed customers seemed very satisfied with the tool, mentioned nothing but praise about the Customer Success-team, and expressed getting the expected value out of their usage.

Lastly, the company who only used the tool twice a week expressed the risk in focusing too narrowly on activity parameters when determining a customer’s health. According to that interviewee, such a parameter could be very much business specific. That customer for example received enough value just by using the tool around one hour a week, for two reasons. 1) Their good understanding of their metrics and 2) their use of numerous other marketing software tools. However, the interviewee noted that if Funnel is the only tool used for tracking a business’ whole advertising performance one could assume that such a company would spend more time in it. According to him, the usage pattern probably wouldn’t just vary between different customer segments, but also after how data-driven the businesses were. Individual differences in how fast new customers get value out of the product was another thing mentioned. According to him a person coming from a business intelligence-background probably would grasp the concept and see value in the tool quicker than a person coming from a marketing background.

5.3 Results from customer web survey

In this section, all of the results from the customer web survey will be presented. The results are divided into quantitative results, stemming from the quantitative questions in the survey, and qualitative results coming from the qualitative question. The following sections: Quantitative results from customer web survey and Results from data mining execution will offer enough results to answer the last research question:
Research question 4:

*What parameters can be used to indicate customer health for Funnel’s customer base and are some parameters more important than others?*

5.3.1 Quantitative results from customer web survey

The web survey was sent to 479 users, from which 84 responses were assembled. Out of these 84 responses, 68 belonged to different companies. For the curious reader, all distributions of each respective survey question can be found in Appendix D. To begin with figure E shows the distribution of the companies’ ratings of Funnel’s customer support, ranging from 1 (very poor) to 5 (very good).

![Support](image)

**Figure E**: Bar chart illustrating the distribution of responses for the support-question in the survey. The white boxes in the bars show the amount of cases as well as the percentage of the whole.

A crosstabulation test between loyalty (*NPS*) and retention (*Likelihood to stay*) was also performed to examine if there existed any relation between these two variables in the survey, see table A. As can be seen, the output shows a difference between expected and real count in the table’s different cells, a difference that was also confirmed by showing statistical significance (p<0.001) in a Chi-Square test, see table B. The dispersion of the answers can be seen in figure F for a more visual illustration. Worth mentioning is that out of the 84 survey answers, only 83 responded on how likely they were to continue being a customer within three months. The only non-response later on churned, having indicated to be satisfied with the product and belonging to the NPS-group “Passives”. One individual also didn’t respond on the NPS-question, resulting in a total of 82 responses in the test below.
Table A: The results from a crosstabulation with the NPS-score as row variable and Likelihood to stay as column variable. Shows the difference in count versus expected count. “Total” shows the total number of responses for each row and column, summing up to 82 responses.

<table>
<thead>
<tr>
<th>NPS</th>
<th>Detractors</th>
<th>Count</th>
<th>Expected Count</th>
<th>Count</th>
<th>Expected Count</th>
<th>Count</th>
<th>Expected Count</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>1.3</td>
<td>7</td>
<td>5.2</td>
<td>2</td>
<td>6.5</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>3</td>
<td>2.0</td>
<td>13</td>
<td>8.5</td>
<td>5</td>
<td>10.5</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Passives</td>
<td>1</td>
<td>4.7</td>
<td>19.3</td>
<td>34</td>
<td>24.0</td>
<td>41.0</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Promoters</td>
<td>1</td>
<td>8.0</td>
<td>33.0</td>
<td>41.0</td>
<td>41.0</td>
<td>82.0</td>
<td></td>
</tr>
</tbody>
</table>

Table B: Chi-square test for the crosstabulation in table A. The test shows statistical significance (p<0.001), see column “Asymptotic Significance (2-sided)”.

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>24,497$^a$</td>
<td>4</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>24,916</td>
<td>4</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>21,535</td>
<td>1</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 3 cells (33.3%) have expected count less than 5. The minimum expected count is 1.27.

Table F: Bar chart illustrating the distribution of the Likelihood to stay-variable between the NPS-segments. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
To further test whether there existed a correlation between any of the survey variables *Satisfaction*, *NPS* and *Likelihood to stay*, a test of variance (ANOVA) was conducted for each relationship. This means that the mean-values for each grouped variable was observed in relation to the chosen dependent variable. It was for example tested how the mean-values of satisfaction differed between the three NPS-groups: “Detractors”, “Passives”, and “Promoters”. All of the tests of variance showed statistical significance for each relationship, with the following results:

- As the satisfaction-level increased, so did the level of loyalty (*NPS*)
- As the satisfaction-level increased, so did the retention level (*Likelihood to stay*)
- As the level of loyalty (*NPS*) increased, so did the retention level (*Likelihood to stay*)
  (the opposite relation to loyalty and retention showed the same pattern)

To study the relationships between the variables in the survey and the parameters from the data mining execution, the survey answers were reduced to only show one response from each company. This was done in order to avoid the issue of double counting companies that had more than one respondent on the survey, resulting in a total of 68 cases. However, it was decided to only study companies that were paying for the use of the web application tool, Dashboards & Reports, resulting in some additional companies being left out. This decision was taken due to the fact that the web application is the only product that is fully monitored and is therefore the only product that shows the complete usage pattern for each customer. Lastly, outliers beyond three standard deviations were also omitted from the data originating from the data mining execution. This was done in order to sort out extreme values that would skew mean values and result in bigger standard deviations. In total, survey answers from 50 companies were studied in relation to the data mining parameters. The parameters that were extracted from the database are shown in table C as well as the descriptive data describing each company. Regarding what parameters to study, parameters were chosen according to how often they had been mentioned during the employee and customer interviews. The parameters shown in table C are therefore the most relevant to study according to the interviewees. A short description of the parameters follows below.

<table>
<thead>
<tr>
<th>Descriptive data:</th>
<th>Parameters to study:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company name</td>
<td>Recency</td>
</tr>
<tr>
<td>Subscription ID</td>
<td>Activity ratio (frequency)</td>
</tr>
<tr>
<td>Business segment</td>
<td>Number of active users</td>
</tr>
<tr>
<td>Subscribed product</td>
<td>Created reports</td>
</tr>
<tr>
<td>Status</td>
<td>Interactions (support tickets)</td>
</tr>
<tr>
<td></td>
<td>No. connected advertising sources types</td>
</tr>
<tr>
<td></td>
<td>No. connected advertising source</td>
</tr>
</tbody>
</table>

Table C: The different data mining parameters that were extracted from Funnel’s database.

**Business segment:** Customers are divided into four segments according to their business type, namely: E-commerce, agency, SaaS and other.
**Subscribed product:** Funnel’s three different products: Dashboards & Reports, the API-connection, and GA Upload.

**Status:** Whether the customer is active or churned.

**Recency:** Amount of days passed since last being active in the tool.

**Activity ratio:** Ratio between amount of active days and amount of days subscribed.

**Number of active users:** Number of active users the last 14 days, where an active user is defined as a user that had at least one activity in the tool during this time period.

**Created reports:** Amount of reports created in the tool, calculated as total number of created reports divided by amount of days subscribed.

**Interactions:** Amount of interactions with Funnel, registered as mail or in app messages. Also called support tickets. Calculated as total number of interactions divided by amount of days subscribed.

**Connected advertising source types:** Amount of different advertising platforms that is connected to a customer’s Funnel-account. For example, Facebook, Google AdWords and LinkedIn (sums up to three source types).

**Connected advertising source:** It is not uncommon for customers to have more than one ad account on each advertising platform, for example three Facebook ad accounts, one Google AdWords account and two LinkedIn accounts (which makes six sources).

To begin with, an ANOVA test was run in order to find out whether there existed a statistically significant difference in the mean values between the observed data parameters and the companies’ likelihood to stay. The results from the test are shown in table D, and showed statistical significance for two parameters: Recency and Activity ratio, see table E. The same test was also run with the NPS as factor, resulting in no statistical significance for any parameter.
Table D: The results from an ANOVA test with Likelihood to stay as factor and the data mining parameters as dependent variables. Column “N” shows the number of responses for each factor-group. Column “Mean” shows the factor-groups mean-values for each parameter. Column “Std. Deviation” shows each row-group’s standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>16</td>
<td>27</td>
<td>49</td>
</tr>
<tr>
<td>Mean</td>
<td>14.50</td>
<td>2.75</td>
<td>2.11</td>
<td>3.84</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>20,157</td>
<td>4,187</td>
<td>3,142</td>
<td>8,333</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>16</td>
<td>27</td>
<td>49</td>
</tr>
<tr>
<td>Mean</td>
<td>.2983</td>
<td>.5294</td>
<td>.5811</td>
<td>.5296</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.20262</td>
<td>.22927</td>
<td>.20618</td>
<td>.22783</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>15</td>
<td>25</td>
<td>46</td>
</tr>
<tr>
<td>Mean</td>
<td>.83</td>
<td>3.00</td>
<td>2.80</td>
<td>2.61</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.753</td>
<td>2.138</td>
<td>1.893</td>
<td>1.972</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>14</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>Mean</td>
<td>.28701</td>
<td>.43059</td>
<td>.32300</td>
<td>.35168</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.181149</td>
<td>.355410</td>
<td>.257112</td>
<td>.283067</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>15</td>
<td>26</td>
<td>47</td>
</tr>
<tr>
<td>Mean</td>
<td>.06291</td>
<td>.08557</td>
<td>.12387</td>
<td>.10386</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.101874</td>
<td>.108301</td>
<td>.221842</td>
<td>.178516</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>14</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>Mean</td>
<td>6.50</td>
<td>8.00</td>
<td>8.48</td>
<td>8.07</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>5.128</td>
<td>3.637</td>
<td>3.653</td>
<td>3.822</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>16</td>
<td>26</td>
<td>48</td>
</tr>
<tr>
<td>Mean</td>
<td>18.67</td>
<td>42.38</td>
<td>37.54</td>
<td>36.79</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>14.638</td>
<td>35.236</td>
<td>31.072</td>
<td>31.392</td>
</tr>
<tr>
<td></td>
<td>Sum of Squares</td>
<td>df</td>
<td>Mean Square</td>
<td>F</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
<td>----</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Recency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>781,527</td>
<td>2</td>
<td>390,764</td>
<td>7,046</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2551,167</td>
<td>46</td>
<td>55,460</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3332,694</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Activity Ratio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>393</td>
<td>2</td>
<td>.196</td>
<td>4,301</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2,099</td>
<td>46</td>
<td>.046</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2,492</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Active users</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>22,123</td>
<td>2</td>
<td>11,062</td>
<td>3,112</td>
</tr>
<tr>
<td>Within Groups</td>
<td>152,833</td>
<td>43</td>
<td>3,554</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>174,957</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.133</td>
<td>2</td>
<td>.066</td>
<td>822</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3,393</td>
<td>42</td>
<td>.081</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3,526</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Created reports</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.025</td>
<td>2</td>
<td>.013</td>
<td>389</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1,440</td>
<td>44</td>
<td>.033</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1,466</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. Source Types</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>19,060</td>
<td>2</td>
<td>9,530</td>
<td>642</td>
</tr>
<tr>
<td>Within Groups</td>
<td>623,740</td>
<td>42</td>
<td>14,851</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>642,800</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. Sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>2484,372</td>
<td>2</td>
<td>1242,186</td>
<td>1,275</td>
</tr>
<tr>
<td>Within Groups</td>
<td>43831,545</td>
<td>45</td>
<td>974,034</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46315,917</td>
<td>47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table E: Corresponding test of significance to the test of variance in table D. The test indicates statistical significance (p<0.005) for the variable Recency and (p<0.05) for Activity ratio, see column “Sig.”

Lastly, a linear stepwise regression analysis was run with the survey answers, one from each company, as the sample. **Likelihood to stay** was chosen as the dependent variable and the same data parameters as above were chosen as independent variables. The results, shown in table F, shows that the data parameter **Activity ratio** succeeds in describing 16.1% of the total variation in the dependent variable **Likelihood to stay** (see adjusted R²-value), with the following model of explanation:

\[
\text{Likelihood to stay} = 3.61 + 1.434(\text{Activity ratio})
\]

Formula A: Linear regression equation with **Likelihood to stay** as dependent variable and **Activity ratio** as independent variable.

**Activity ratio** was in other words the only parameter proved to correlate with the **Likelihood to stay** in the regression model. The same analysis was run with the **NPS-score** as dependent variable, but with no parameters entered into the model of explanation.
Table F: The results from a linear stepwise regression analysis with Likelihood to stay as dependent variable and the data mining parameters as independent variables. The first table shows the only variable that succeeded in entering the model: Activity ratio. The second table shows the Model summary with an adjusted $R^2$ value at 0.16. The third table shows the ANOVA test indicating statistical significance ($p<0.05$), see column “Sig.”. The “Coefficients”-table shows the values that build up the final regression-model, see column “Unstandardized B”.

### Variables Entered/Removed

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables Entered</th>
<th>Variables Removed</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Activity Ratio</td>
<td></td>
<td>Stepwise (Criteria: Probability-of-F-to-enter $\leq$, 050. Probability-of-F-to-remove $\geq$, 100).</td>
</tr>
</tbody>
</table>

- a. Dependent Variable: Likelihood to stay

### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.426</td>
<td>.181</td>
<td>.161</td>
<td>.725</td>
</tr>
</tbody>
</table>

- a. Predictors: (Constant), Activity Ratio
- b. Dependent Variable: Likelihood to stay

### ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4,643</td>
<td>1</td>
<td>4,643</td>
<td>8.844</td>
<td>.005</td>
</tr>
<tr>
<td>Residual</td>
<td>21,000</td>
<td>40</td>
<td>.525</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25,643</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- a. Dependent Variable: Likelihood to stay
- b. Predictors: (Constant), Activity Ratio

### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>3,610</td>
<td>.275</td>
<td>13.120</td>
<td>.000</td>
</tr>
<tr>
<td>Activity Ratio</td>
<td>1,434</td>
<td>.482</td>
<td>.426</td>
<td>2.974</td>
</tr>
</tbody>
</table>

- a. Dependent Variable: Likelihood to stay

5.3.2 Qualitative results from customer web survey

The qualitative results from the survey stem from a follow-up question on the quantitative question: “How likely are you to still be a customer in three months?” The idea of including such a question was to let respondents elaborate on their previous answer, which would give a better understanding of their rating. In total 61 text answers were received. The general observation that could be made was that for those respondents who had responded with a rating of “Fairly likely” or “Extremely likely”, the text answers included mentions of the
superiority of the tool compared to competitors, the product and service working as promised, the functionality, dependence on the tool, as well as lock-in due to contractual or technical aspects. The mid-section of the answers in the range “Unsure” to “Fairly likely” included mentions of how the decision depended on other persons or factors, that the respondent was looking at other options, and how the decision depended on how the tool developed. Lastly, three respondents responded with “Fairly unlikely” and could easily be grouped into one cluster. These answers all mentioned problems with the product or set up, which had not been resolved as expected by the customer. The complete summary of all the answers as well as their clusters can be found in appendix E.

5.4 Results from data mining execution
To begin with, data on 278 customers were extracted from Funnel’s database, belonging to customers using the web application Dashboards & Reports. Additionally, outliers were detected as data points three standard deviations away from each parameters’ mean value and omitted from the material.

In order to examine if any of the data parameters had any relation to whether a customer churned or not, all of the parameters were added as dependent variables into an ANOVA test with the variable Status as factor. The results are shown in table G and show statistical significance (p<0.001) for all of the parameters except Interactions and Created reports, see table H.
Table G: The results from an ANOVA test with Status as factor and the data mining parameters as dependent variables. Column “N” shows the number of responses for each factor-group. Column “Mean” shows the factor-groups mean-values for each parameter. Column “Std. Deviation” shows each factor-group’s standard deviation.
Table H: Corresponding test of significance to the test of variance in table G. The test indicates statistical significance (p<0.001) for all variables except Interactions and Created reports, see column “Sig.”

To follow up on this and test the correlation between the parameters and the customers’ status, a logistic regression analysis was conducted with Status as dependent variable. To begin with, a test of linearity was performed in order to check whether a linear relation existed between the abovementioned parameters, see appendix F. This is a prerequisite that must be met for each parameter included in the test, in order to get trustworthy results. From this examination linearity was shown between all parameters except Recency and Created Reports, for this reason these two parameters were excluded from the rest of the analysis.

After that, all of the remaining parameters were set as independent variables and the regression analysis was conducted. However, the logistic regression was repeated in order to see how the exclusion of parameters not showing statistical significance could improve the explanation model. The logistic regression was in other words iterated until the best explanation model had been found, resulting in keeping two variables: Activity ratio and Number of source types, which both showed statistical significance (p<0.001) and resulted in the following explanation model, see table I:

\[
Status = -2.131 + 3.048(\text{Activity ratio}) + 0.351(\text{Number of source types})
\]

Formula B: Logistic regression equation with Status as dependent variable and Activity ratio and Number of source types as independent variables.
Lastly, the classification table, see table J, shows how well the model succeeded in predicting the customers as either churned or active. As can be seen in the table, the model succeeded in predicting the correct status of 74.5 % of all customers, compared to the null model which only succeeded in predicting 56.8 %, see table K.

<table>
<thead>
<tr>
<th>Variables in the Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Step 1 Act. Ratio</td>
</tr>
<tr>
<td>No. Source Types</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

a. Variable(s) entered on step 1: Activity Ratio, No. Source Types.

Table I: The results from a binary logistic regression analysis with Status as dependent variable and the data mining parameters as independent variables. Two variables succeeded in entering the model: Activity ratio and Number of source types, showing statistical significance (p<0.001), see column “Sig.”. The values that build up the final regression function are seen under column “B”.

**Classification Table**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted</th>
<th>Percentage Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Churned</td>
<td>Active</td>
</tr>
<tr>
<td>Step 1</td>
<td>84</td>
<td>36</td>
</tr>
<tr>
<td>Active</td>
<td>35</td>
<td>123</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. The cut value is .500

Table J: The classification table showing how well the logistic model succeeded in defining the cases as either active or churned, with an overall predictive ability of 74.5 %.

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted</th>
<th>Percentage Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Churned</td>
<td>Active</td>
</tr>
<tr>
<td>Step 0</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>Active</td>
<td>0</td>
<td>158</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Constant is included in the model.
b. The cut value is .500

Table K: The null model, with an overall predictive ability of 56.8 %.
6 Analysis
In this section, the results will be analyzed and the research questions that was stated in the beginning of the report will be answered. For the ease of reading, the chapter will be divided into two parts, the first dealing with the definition of customer health and how it relates to customer satisfaction, loyalty and retention. The second part will go into greater detail in what parameters that might be used to track customer health at the studied company Funnel.

6.1 Customer health and the relationship between satisfaction, loyalty and retention
The very first thing to consider when starting this study was how to define customer health, as was stated in research question 1:

What is customer health and how can previous theory and literature help us understand this concept further?

But also, to understand what customer health actually means for the employees at Funnel, the people that are going to use this metric to try and measure their customers’ well-being. Therefore, research question 2 was stated as follow:

What is seen as good and poor customer health at Funnel and how can this be identified in customers’ product usage?

After consulting previous theory and literature two things became obvious 1) the literature on this concept is very recently published with no mention in the academic field and 2) there exist no clear definition of what this concept actually tries to capture. The definition by Forrester Consulting (2014) mentions the correlation to churn and retention, which was also mentioned by Funnel’s employees as well as customers. However, other definitions include concepts such as customer satisfaction, loyalty, and perceived value (Raboy, 2013; Abbott, 2017). After examining the SPC-litterature, it is clear that the linkages between perceived value, satisfaction, loyalty, retention, and profitability are all the more studied and could be used to understand customer health further.

Looking at what Funnel’s employees would classify as good or poor customer health, the propensity for retention or churn was mentioned as one important aspect. Another element to consider was the customers’ likelihood to be an advocate for the firm or spread bad word of mouth. According to Frederick Reichheld (2003), customer recommendations are one of the best indications of customer loyalty and retention, and following this logic a good customer health should indicate a loyal customer that stays with the company while a poor customer health should indicate disloyalty and high propensity for defection. It is surely not a coincidence that Funnel’s employees describe customer health in this way, since the NPS-score, developed by Frederick Reichheld, is widely used in the SaaS industry and is assumed to signal the true loyal customers, while filtering out the ones that stay because of inertia, indifference and exit barriers (Reichheld, 2003).
To test this proposition and examine if it would hold true for Funnel’s own customer base a web survey was sent out to the customers. The results after looking at the customers NPS-score in relation to their indicated propensity to stay with the company shows a somewhat ambiguous result though. Starting with the crosstabulation, see table A, the test shows an overrepresentation of respondents “Extremely likely to stay” in the NPS-group called “Promoters”, as well as less counts than would be expected in the other two groups. Respondents answering “Fairly unlikely/ Unsure” and “Fairly likely” are on the other hand underrepresented in the “Promoters”-group, but overrepresented in the “Detractors”- and “Passives”-group. This goes in line with the additional test of variance that was conducted on the same relationship, showing that as the NPS-score increased, the likelihood to stay did so as well. In other words, the general tendency in the results gives support for the SPC-model, see figure A, and the relationship between loyalty and retention.

However, by examining table A more closely as well as the bar chart depicting the same distribution of answers, see figure F, several cases can be found contradictive to this general trend. To begin with, the “Passives”-group includes a good mix of customer’s both indicating that they are going to stay but also more unsure customers. This makes this whole group hard to interpret and one should be careful when drawing any conclusions about this group’s general retention-level. The “Detractors”-group on the other hand includes a few customers indicating both a “Fairly likely” chance as well as an “Extremely likely” chance to stay as customers. This could indicate customers that stay because of reasons such as inertia or exit barriers as mentioned by Reichheld (2003). What is more interesting though is customers that seem to work as advocates for the company, still seem unsure whether they are going to stay or not. The two customers indicating such attitudes, see the “Promoters”-groups in figure F, were followed up internally to cast more light on the situation. One of these customers had recently decreased their marketing spend due to financial circumstances, hence got less value out of Funnel relative the cost paid for the tool. The other respondent used the web tool through an agency and had experienced syncing issues with the tool. However, after quit using the tool through their agency, the company had signed up for an own subscription and is at the time of writing still using the tool. At first glance, these two cases seem contradictive to the general idea that loyalty leads to retention. However, after looking into it we can conclude that external factors are the reason to why these two cases stand out from the rest. This is probably the take away from this more granular investigation. Although the general tendency shows a positive correlation between loyalty (NPS-score) and retention (likelihood to stay), external factors often influence in real case scenarios, which leads to deviations from the general theory.

Rewinding back to research question one, Funnel’s employees mentioned perceived value as a good definition of customer health as well as the quality of the relationship between company and customer. In the Theory and Literature-chapter, studies by Yang and Peterson (2004), Mittal (2000), and Sirdeshmukh et al. (2002) showed positive correlations between perceived value and customer satisfaction and loyalty. In addition, relationship quality was found to have a mediating role in the SPC-model by Kumar & Petersen (2012, p.63). However, to investigate perceived value’s relation to customer health further and look closer
at the relationship quality between company and its customers, customer interviews were conducted and a web survey was sent out. This relates to research question 3:

What value is the product providing its users and how is this related to customer health?

To begin with, all interviewed customers expressed getting the expected value out of the tool, hence the perceived value seemed to be high for these customers. Regarding the satisfaction-level, the general feeling from the interviews was that every interviewee seemed very satisfied with their subscription, however such a hunch should not be used to draw any conclusions from. In that case, it is better to refer to their survey answers, in which all stated to be satisfied with the tool as well as showing good levels of loyalty on the NPS-score and expressing a very high chance of staying as customers. Worth mentioning here is the benefits that can be gained by surpassing a customer’s expectations on a product or service. One of the interviewees recalled being satisfied with the value provided by the tool, however when introduced to a new functionality this turned out to create additional value for the customer. One year later the customer’s business is partly dependent on this functionality. One can assume, by looking at the literature on perceived value and satisfaction by Anderson and Mittal (2000) and Walter, Thilo, and Helfert (2002), that such an experience results in high levels of satisfaction as well as creates the best condition for long-term loyalty and retention.

Concerning the relationship quality, the customers that were interviewed seemed to be very satisfied with their contact with the company and expressed having received good support. This was further studied in the web-survey, resulting in very good scores for the support, see figure E. This indicated that Funnel’s customer relations seemed to be of good character, hence this relationship was not further examined. However, since relationship quality has been found to have a mediating role between customer satisfaction and retention and loyalty in the SPC-model, see figure A, it is possible that this is also the case in Funnel’s customer relations. To summarize, the earlier proposed linkage between a high perceived value and high satisfaction- and loyalty-levels seemed to hold true for the customers interviewed, and the relationships was indicated to be of good quality.

At this point we might have found what the concept customer health actually tries to capture. If a customer’s perceived value is on the same level as his or her expectations, the customer will be satisfied with the purchase, hence continue to do business with the company as long as no better option appear, all according to equity theory (Bolton & Lemon, 1999) and research on perceived value (Yang & Peterson, 2004). If the customer’s perceived value exceeds the expectation though, the customer will be overly satisfied with the purchase, hence more likely develop a stronger affective loyalty towards the brand, which makes the chance of the customer staying even stronger. All in line with research by Allen (2004). However, if the customer experiences issues with the product or service, the perceived value might turn out to be less than the expected value, leading to dissatisfaction with the tool and a decrease in loyalty. Since literature by Kumar and Reinartz (2012) has shown that dissatisfaction has a stronger effect on retention than satisfaction, see figure B, such a customer will most certainly
be in the risk zone of churning. That is if no efforts are made to deliver the value that the customer expects in time. Customer health should therefore be defined as the outcome and measure of a customer’s perceived value. This can be observed in the qualitative survey answers where all customers that reported low retention-levels had experienced problems with their subscription. The relation between expected and perceived value is well illustrated by the following quote from such a customer response coming from Appendix E:

[Experienced] issues with data sync that weren't resolved […], didn't have time to keep troubleshooting and stopped relying on the data. Planning to cancel, as I don't have time to figure it out. LOVE the concept and wish it were more affordable so I could use it for all of my clients, otherwise it makes more sense to manually upload the data to analytics each month. – Customer response on web survey, 2017-03-28

To clarify, other definitions of customer health that includes the level of satisfaction, loyalty and retention are still valid. However, these concepts are just additional consequences of a customer’s perceived value, and fail in describing the whole picture. Concerning the influence of relationship quality on customer health, this is an aspect that has not been studied closer. However, since the relationship quality has been found to have a mediating role in the satisfaction-loyalty-profit chain, it is likely that the relationship quality between company and customer also affects a customer’s health.

Finally, it should be mentioned that a high perceived value, measured as good customer health, not certainly equals customer retention. External factors such as how the decision to continue lies on someone else in the organization, or financial constraints, can all lead to a churn even though the customer health is good. When we now have a definition of customer health, as well as have understood how the underlying concepts relate to each other, we can move on to look at how perceived value, or customer health, correlates with product usage.

**6.2 Parameters indicating customer health**

This section will go over the analyses that have been conducted on the data mining parameters, as well as the survey data, and aims at answering research question 4:

"What parameters can be used to indicate customer health for Funnel’s customer base and are some parameters more important than others?"

**6.2.1 Likelihood to stay and NPS versus data mining parameters**

Beginning with the test of variance (ANOVA) comparing the data mining parameters and the survey’s retention-variable Likelihood to stay, see table D, the test succeeded in showing statistical significance for two variables: Recency and Activity ratio, see table E. What can be read from this is that users that used the tool often indicated a higher propensity to stay as customers, than customers that used the tool more seldom. However, it should be noted that the standard deviation for the parameter Recency differs greatly between the factor-groups, so this can have resulted in the parameter showing statistical significance (p<0.005). After all, Activity ratio is the only variable that enters the subsequent linear regression model, see table F, so the relationship between this parameter and the likelihood to stay is stronger than for
Recency. It should also be noted, that even though the rest of the studied data mining parameters didn’t reach statistical significance (p<0.05), all of them showed an increase in mean-values between customers that are “Fairly unlikely/unsure to stay” and “Fairly or Extremely likely to stay”. It is therefore possible that more than two parameters would have shown statistical significance if the sample had been bigger. As mentioned in the Empirical results-section, the same test was run but with the NPS-score as factor variable. This test failed in showing statistical significance (p<0.05) for any of the parameters, which means that no correlation could be found between the parameters and the different loyalty-levels.

6.2.2 Status-variable versus data mining parameters

As have been mentioned, good customer health doesn’t have to imply retention. However, when looking at the bigger sample of Funnel’s active and churned customers, churned customers were seen as an outcome of poor customer health, while active customers were seen as being of better health. This assumption can be questioned given the fact that external factors have been shown to affect a customer’s decision to churn or not. However, as have been mentioned in both literature and previous reasoning, the general notion is that high perceived value correlates more closely with retention than low perceived value. That being said, active customers were in the following analysis seen as having experienced a higher perceived value than the churned customers, resulting in active customers being of better customer health than churned.

Beginning with the ANOVA test, this test found a statistical difference in mean values between churned and active customers for the following parameters: Recency, Activity ratio, Number of active users, Number of source types, and Number of sources, see table G and H. For the sake of clarity, this section will go through the results more closely. We can for example see that the Recency-parameter, the amount of days since the last event in the tool, is far greater for churned customers than for active users. Activity ratio shows that churned customers are on average spending about 17 % of their full lifetime in the tool, while active customers are using the tool about 39 % of all their days subscribed. Number of active users, defined as the number of users that have been active in the tool at least once the last 14 days, is also higher for active customers than for churned. Lastly, active customers have in general about the double number of sources and source types compared to churned customers.

The difference in Number of source types and Number of sources for churned and active customers could be interpreted as the more advertising platforms and ad accounts you use, the bigger problem is the tool solving for you, and therefore the more perceived value you experience as a customer. For example, if a customer signs up for a subscription and only connects a few advertising platforms, the customer might eventually realize that the problem wasn’t as big as imagined and might opt out and try to manage the problem themselves. In such a scenario, the expected value was greater than the perceived value, which resulted in a churn. If the customer on the other hand manages several ad platforms and ad accounts, the time saved might be well worth the money.
A logistic regression was also performed in this study to see if there existed any relationship between any of the data mining parameters (independent variables) and the Status of the customers (dependent variable). The results showed that the final model could predict the correct status of 75% of the customers, see table J, by looking at the two variables Activity ratio and Number of source types, see table I, indicating that these two parameters give fairly good means of determining whether a customer will churn or not. Hence, when studying all parameters, these two might be the most important when measuring customer health.

Looking at all parameters all together, Activity ratio is the one that stands out as the one variable that correlates with the compared variables in all statistical analyses so it should be clear by now that this parameter is the most important of all studied. However, what is maybe more interesting when looking at all parameters is that a higher product usage seems to correlate positively with the propensity to stay as a customer. This is both seen when comparing churned and active customers, as well as looking at their self-reported likelihood to stay. In practice, this might be interpreted as the more you use a product, the more perceived value you experience, hence this leads to customer retention and a good customer health. However, the observant reader might sound a note of warning considering the fact that good customer health at this point seems to be equated with a high activity level. This is a valid point considering the fact that several of the interviewees mentioned that a low activity not necessarily must imply that a company isn’t getting their expected value out of the tool and that a high activity also could indicate a struggling customer. However, although this might hold true for a few customers, the general tendency illustrated by the statistical analyses seem to indicate that a high product usage more closely correlates with good customer health, rather than poor. One way to get around this issue would be to investigate what kind of customer that you are dealing with during set-up in terms of professional background, previous knowledge of marketing practices, use of complementary marketing softwares, and the company’s general need to follow up on their marketing efforts. There are surely other aspects that could be considered. However, that aside, a high product usage will work as a signal of good customer health for most customers.

6.3 Critical evaluation of results and analysis
We will end this section with a critical evaluation of the results and analysis. To begin with, there is a risk that the data from the data mining execution have been prone to several errors. Apart from being the object of human errors as indicated in the Method and execution-section, some of the results heavily depend on how each parameter has been defined when extracting the data. All the time dependent variables such as Recency, Activity ratio and Number of active users should be viewed with care when looking at churned customers. The reason for this is that the churn date for a customer has been defined as the last day having status ‘active’ in the billing system. In other words, even though a customer calls in and end their subscription before the last day in their billing cycle, the customer is going to be marked as ‘active’ in the billing system until the last day has passed, when it will change to ‘churned’. This means that even though a decision has been taken to end the subscription this won’t show until the customer’s last paying date has passed. This have implications for the time-dependent variables in such a way that the days of inactivity between the churn-decision and
the status-change in the billing system will portray these customers as more inactive than they might have been before the actual churn decision. However, it is impossible to know for sure when a decision to end a subscription is taken. This would require the status-information to be extracted from a data table where incoming information from the customer is stored. Such a data table doesn’t exist. In other words, the means for determining a customer’s “real” churn date are very limited. However, the difference in these parameters mean-values between churned and active users might have been amplified for this reason.

That being said, we can conclude from the analysis between the data mining parameters and Likelihood to stay-variable, see table D, that the mean-values for the time-dependent parameters also increases with their stated propensity to stay. Since all the surveyed customers were active at the time of data extraction, we can conclude that our recently mentioned doubt regarding the accuracy of data might not be as grave as feared.

A dubiety concerning the quantitative data coming from the web survey is that this data has been self-reported. In other words, parameters such as the level of Satisfaction, NPS and Likelihood to stay are based on the respondents’ own self-estimations. This can be questioned for two reasons, 1) self-estimations very much depend on the individual answering the questionnaire and could be affected by both current mood and personal characteristics, and 2) the correspondence between self-assessments and real behavior have been questioned by earlier mentioned literature (Trout & Rivkin, 2008).

Lastly, it should be said that the results on which this analysis has been conducted is all based on the data that has been extracted at the time of the data mining execution. This means that as the customer base is changing, as well as the product, these conclusions might become invalid and turn the findings inaccurate. In that sense, the reliability of the study is dependent on the data that is used as input.
7 Discussion and conclusions

This last concluding section will wrap up the study by referring back to the background for this thesis work as well as the purpose that was stated in the beginning of the report. The chapter will start by discussing theoretical implications of this study, and then continue to go through practical implications. Thereafter the report will be concluded and suggestions for future research will be covered.

7.1 Theoretical implications

This report began by mentioning the shift towards more long-term business relationships that a lot of business fields are experiencing today. This has led to the emergence of business approaches such as Customer Relationship Management (CRM) and Customer Success, which are focused on maximizing the value for both customer and company. However, while CRM has been around for a while and has had the time to define activities and processes as well as developed its own terminology, Customer Success seems to experience growing-pains. This has become apparent in this study when investigating the new concept customer health. What is remarkable is that existing literature about Customer Success stresses the importance of measuring this new concept by the use of customer health scores, yet no solid definition of what customer health actually means can be found. This could possibly be due to the fact that Customer Success stems from the business field, rather than the academic field. Generally speaking, these two fields differ in such a way that concepts coming from the academic field are the outcome of research and have been object to great scrutiny, while concepts from the business field seldom have to go through the same tough process in order to be accepted as valid. As have been mentioned in the Method and execution chapter, validity in the academic field is reached when research succeeds in showing results, on which coherent conclusions can be drawn (Bryman, 2002, p.50). Questions such as: “Are we measuring what we think we are measuring?” and “Do the proposed causal relationships hold water?”, are therefore of great concern when defining a new concept in the academic field. The Customer Success-literature, however, seems to view these questions as less important and is generally more concerned in delivering cures and solutions on how to handle your customers in the new business climate. The important thing is not the concepts themselves, they can be both vague and fairly fluent, but instead what they are indicating as important for a business. Measuring customer health for example helps the company focus on the individual customer, and offers a foundation on which interactions with the customer can begin.

Furthermore, the definitions of customer health that was mentioned by Raboy (2013), Abbott (2017) and Forrester Consulting (2014) all mentioned retention and churn as important outcomes of good and bad customer health. Such a definition tells more about the implication of a concept, rather than the concept itself. This also indicates that customer health is more of a practitioner’s concept that is focused on easing businesses’ practices, rather than a concept that is the outcome of research and theoretical discussion. What this report can conclude is that perceived value is a good definition of customer health. This matches well with the goal of Customer Success-management, which is to ensure that the customer is getting maximum...
value from his or her purchase (Smilansky, 2016). In other words, if retention or churn is the outcome of customer health, then perceived value is the determinant.

The definition of customer health is only part of this study’s contribution though. Maybe more important is this study’s illustration of the underlying concepts in the satisfaction-loyalty-profit chain. Because even though the field of Customer Success will be able to develop its own business practices without this knowledge, one shouldn’t underestimate the value in grasping the bigger picture. The general theory as described by Kumar and Petersen (2012), together with more granular studies on the chain’s different sections such as the work on customer satisfaction and profit by Anderson, Fornell and Lehmann (1994) and the research on customer loyalty and retention by Allen (2004), have in this study been necessary to understand the concepts’ interconnectedness and draw any conclusions about customer health. However, as also should be clear to the reader at this point, there exists a great deal of contradictive research in these areas. While one can argue that these contradictive studies undermine the just mentioned research, one can also look at them as complementing the already existing theories, making them more complete by adding new knowledge. The complexity of these concepts shouldn’t come as a surprise though. Business research is conducted in the social sphere, which in contrast to natural scientific research involves difficulties with both measurement and context-dependency, just to name a few. The same goes for this study, even though the results might be very much context and company-specific, the aim of studies like this is not to reach general theory but to describe the case at hand with the help of inductive reasoning.

7.2 Practical implications

A more practical contribution of this study is the outlined process of how to find parameters that can indicate a customer’s well-being, the so-called customer health parameters. By glancing at the CRM-practice of developing churn-prediction models this study has been able to develop its own method for how to identify relevant customer health parameters. The method of logistic regression has been used, which together with neural networks and random forests is one of the most commonly used churn-prediction methods (Kumar & Petersen, 2012, p.152). However, while the work surrounding the development and implementation of a churn-prediction model is a heavy process as illustrated in Theory and literature in figure D, the process used in this study has been somewhat easier. The methods that have been presented here can in other words work as a good middle way for smaller companies working with Customer Success. Companies that on the one hand don’t have the time or resources to develop an overly complex churn-prediction model, on the other hand want some more hard evidence than the trial and error-parameters that is suggested by most customer health score vendors. In addition to the process, the customer health parameters that are the result of this study can work as a starting point when similar businesses want to investigate their own customer health parameters. After all, most of the parameters that have been found to be of importance are of the recency, frequency, and monetary (RFM)-character, variables that often appear in churn prediction models. Monetary variables, however, have been excluded from this study, and will be discussed in the last section discussing future research. The studied parameters that are of the recency- and frequency-type however, shows that a higher product
usage most of the times signals good customer health. This goes in line with previous research by Bolton and Lemon (1999) and Ram and Jung (1991) who has showed that satisfied customers are more likely to have a higher service usage level than unsatisfied customers.

So how might a software company use this information to minimize its churn and increase their customers’ well-being? Well, according to this study the health of the customer increases with the software usage, and a good idea is therefore to aim at increasing the customers’ interaction with the product. This can be done, firstly by improving the tool’s functionality so that the tool solves its customers’ problems to a greater extent. Secondly, by making sure that the customers understand all of the tool’s functions and how they can be used to ease the customers’ business. And thirdly, by improving sharing features so that the tool can be spread within the customers’ own organizations. An additional benefit with having several people in an organization using the tool is that this can work as a buffer or safety net when unexpected events occur. Let’s illustrate with an example. If only one person in an organization uses the tool and then quit their job or change position within the company, the software provider loses its only contact at that company. In such a situation, there is a big risk that the following person gets limited, if any, information about the tool, therefore not see the necessity in using it, and therefore end their contract with the provider. If the tool on the other hand would have been used by several people in the organization, these people would probably have continued using the tool regardless of their colleague quitting, and would also have had the possibility to instruct the new person in how it is being used in the organization.

While parameters such as Recency, Activity ratio and Number of active users are linked to the product usage and make sense to study for most software providers, more company specific parameters have also been demonstrated to be of importance when determining a customer’s health. In this study, the Number of source types and Number of sources that a customer connects has been shown to correlate with the propensity to churn or not. For Funnel, this information could be used not only to indicate at-risk customers, but also to determine the fit of a new customer. If the number of ad platforms are very low for example, the customer might not be profitable to acquire in the first place since the customer pose a risk of churning before bringing in enough money to cover for the acquisition costs. The two parameters can also be part of a continuously updated health score since there is a possibility that a customer start using new ad platforms or make changes to their existing marketing practices. A sudden decrease in number of ad accounts should for example raise a red flag for Funnel and indicate a reason to reach out and investigate what is happening with the customer and possibly prevent them from churning.

The practical conclusion that can be drawn from this is that a company can yield several benefits by investigating data not only linked to the product usage, but also data linked to the set-up and the customers themselves. Similar to the use of CRM-systems, in which the goal is to segment a company’s customers after their similar characteristics or needs and develop strategies for how to interact with them (Kumar & Reinartz, 2012, p.19), the process of examining customer data when determining customer health parameters can result in insights that can be used for strategic purposes. In the example above, parameters linked to the
customer’s own business can for example be used to segment customers and make informed decisions on what potential customers to acquire. This will have implications for parts of the organization working with lead-generation and sales, as well as might influence how the company is marketed. In that sense, the exploration of a company’s customer data and creation of a customer health score can not only provide valuable information to the Customer Success team, but to the whole organization.

The bottom line of this study is that a company that is collecting and storing data about its customers can achieve great customer insight by investigating this data. Information about what customer behavior that might forego a churn, for example, is hidden in this data and can have economic implications for a company if discovered in time. The discovery of customer patterns and behaviors also allow the company to segment its customer base better, and open up for more targeted actions that can save the company’s resources. Such possibilities might also improve the customer relationships, since more precise targeting allow the company to tailor their message to each individual customer, which in turn might result in a greater generation of good word of mouth. To summarize, the exploration of data linked to both customers and product usage can benefit a company in many ways. The process outlined in this report is just one out of many that a company can use to reach better customer insight.

7.3 Conclusions

This study set out to explore the newly emerged management philosophy Customer Success, which mission is focused on knowing and understanding the customer’s needs in order to deliver as much value to the customer as possible. By satisfying the customer’s expectations the idea is that this will lead to maximal recurring profit for the company. One way to follow up on how a company’s customer base is doing is by the use of customer health scores. However, since it was apparent from the start that there exist little to no consensus on what customer health actually means, or how to find parameters that can measure this concept, the study’s purpose was:

*To extend the existing knowledge of the business concept customer health and show how to identify relevant parameters for measuring customer health.*

By investigating how the customer health concept is used in the business field strong connections to concepts such as perceived value, customer satisfaction, loyalty, and retention could be found. After a thorough examination, this study can conclude that customer health should be defined as a measure of a customer’s perceived value of a product or service. Additionally, both this study and previous theory have shown that customer satisfaction, loyalty, and retention are all outcomes of a customer’s perceived value. For the sake of clarity though, they should not be used to define customer health. The relationship quality between company and customer might also have a mediating role on customer health, however this can neither be confirmed nor refuted from this study’s findings.

A more granular examination of the level of loyalty and retention among the studied company’s customers showed that the two concepts correlate fairly positively, following the
general line of thought in previous theory and literature. However, retention doesn’t imply loyalty at all times, as well as loyalty doesn’t have to indicate retention. Additionally, loyalty is a long-term outcome of perceived value, while external factors can influence the decision to stay as a customer, even though the perceived value is high. For that reason, good customer health won’t necessary equal retention or loyalty. However, external factors aside, retention was seen as an outcome of high perceived value when the relationship between customer health and product usage later on was studied. From this examination, it could be concluded that a customer’s health generally increases with the product usage. Indicating that perceived value and product usage correlates positively.

Regarding more company-specific discoveries, this study can conclude that the following parameters can be used to track customer health for Funnel’s customers:

- Recency
- Activity ratio
- Number of active users
- Number of source types
- Number of sources

Out of these parameters Activity ratio and Number of source types showed to correlate the strongest with customer health, hence they are of greater importance to monitor and might be weighted heavier in a future customer health score. The activity parameters Recency, Activity ratio, and Number of active users showed that a higher product usage correlate with better customer health and a higher chance of staying as a customer. Such a pattern can be interpreted as the more you use the tool, the more value you get out of it. Hence, if the product usage is low, a Customer Success representative should reach out to the customer in order to see if the customer is doing fine or need help. Regarding the Number of source types and Number of sources that a customer connects to the product, these two parameters can be observed both during the sales process as well as when the customer has started to use the tool. High numbers on these parameters was found to correlate with a higher chance of staying as a customer, hence this can be interpreted as the more ad platforms and accounts that you connect with the product, the bigger problem is the product solving for you.

At this point, the concept customer health has been defined and a process for how to identify relevant customer health parameters has been outlined, hence it can be confirmed that the purpose of this study has been achieved. However, the results are very much dependent on the data that have been studied as well as the processing of data. In that sense, the reliability of the study can be questioned. Because although the measures that have been used have been consistent and followed specified rules, such as the elimination of outliers outside three standards deviations, the results are still all dependent on the basic data that have been studied. However, if replicating the study with fairly similar basic data, similar results would most likely be obtained.

57
Lastly, this study will hopefully add knowledge to the field of Customer Success as well as work as a first attempt to introduce the management philosophy into the academic literature. This study further contributes with results to the literature on satisfaction, loyalty and retention, and links these concepts to the new concept customer health.

7.4 Further research suggestions
As was mentioned in the very beginning of this report, metrics such as customer lifetime value (CLV), customer acquisition cost (CAC) and customer retention cost (CRC) are commonly used in the field of CRM. According to Saleh (2015) 76% of companies view CLV as an important metric to track for their business, however only 42% are able to calculate it correctly. In this study, these monetary variables have not been included when studying Funnel’s customer base. If they were to be added though they could add another dimension for measuring customer health. Adding parameters like these to a customer health score opens up opportunities to indicate both customer health and customer value. Put together this would give the company a complete picture over what profitable customers that needs more attendance, at the same time indicate what customers that consume plenty of resources but yield little in return. For a Customer Success-team this would present good ways of prioritizing what customers with poor health that are the most profitable to retain and therefore should be given immediate attendance.

Another suggestion for future research is to delve deeper into individual product usage styles, since the positive correlation between product activity and customer health might not hold true for all customers. The ultimate software tool would include features that could indicate customer satisfaction or frustration just by monitoring the way a user uses the tool, for example by registering the movement of the mouse. However, this would require a thorough study of a customer’s product usage by for example observing individuals use the tool and let them clarify their feelings and perceptions in connection to this usage. Even though this could result in the best customer insight possible, the time, costs and effort linked to such a study could be substantial.

In that case it is probably a better idea to zoom out a bit and study whether product usage differ among customer segments. The idea that different customer segments might use the tool in different ways have been a recurring suggestion during this study, however the segments that have been studied have been too small to draw any wider conclusions from. This study has therefore been performed at a more general level. That being said, the possibilities to distinguish such differences in product usage increase with the size of a company’s customer base, so as a company grows by allocating resources to the right customers, so does its opportunities to fine-tune the ways of measuring customer health.
8 References


9 Appendices

9.1 Appendix A: Interview template for interviews with Funnel’s employees

Bakgrund
- Namn
- Yrkesroll
- Ansvarsområde

Allmänt kundarbete
- Vad gör Customer Success-teamet idag?
- Vad kan Customer Success inte göra idag?
  o Hur läser ni av kundernas “mående” idag?
- Hur ser kontakten mellan företag och kund ut?
  o Varierar den för respektive kund?

Customer health
- Vad skulle du säga att customer health är?
  o Hur bör man definiera ett gott vs dåligt mående?
- Vilka parametrar tror du bäst indikerar customer health för en kund?
  o Kan någon av dessa parametrar vara missvisande?

Customer health score
- Vad bör ett customer health score kunna indikera?
- Vilka parametrar bör ingå i ett sådant Health Score?
  o Tror du alla parametrar är lika viktiga eller bör de viktas?
- På vilket sätt skulle ett customer health score kunna påverka/underlätta verksamheten?

Användarmönster
- Idag kan vi ju studera Heat Maps för att få en uppfattning om kundernas produktanvändning. Vilka för och nackdelar skulle du se med att studera dessa användningsparametrar för att mäta customer health?
  o Recency- Hur lång tid som passerat sen senaste användandet
  o Frequency- Hur ofta produkten används inom en given period
  o Activity- Antalet aktiviteter per dag
  o Depth of usage- användande i % av verktygets funktioner
- Brukar användarmönstret förändras över kundcykeln?
  o Hur i så fall? Finns det någon generell förändring? Några tidsintervall?
- Skiljer sig användarmönstret åt mellan olika kundsegment?
- Finns det några sticky features som man visar på ett större engagemang/bättre mående från kundens sida?

Last Question
- Hur skulle du lösa problemet?
  o Vilka kunder skulle du intervju? 

Finns det något som jag missat eller bör tilläggas?
9.2 Appendix B: Interview template for interviews with Funnel’s customers

Background
- Name
- Job title
- Field of responsibility

Background
- Describe shortly what your company is doing and your role in the firm.
- How did you get in touch with Funnel the first time?
- How many use Funnel at your company and what for?
- How many at your company come in contact with the tool? (In case more people get in contact with the tool’s functions such as reports created using Funnel or similar)
- How often do you use the tool?

Usage
- Why are you using Funnel?
- What are you using Funnel for?
- Which functions are the most important for you (key functions)?

Product value
- Can you elaborate on the value that the tool gives you?
- Have you received the value that you expected from the tool?
- What are your thoughts on the tool qualitywise?
- Is it easy to use? Intuitive?
- Is it working as expected?
- Is there any function that you miss in Funnel?

Company contact
- How often are you in touch with Funnel’s Customer Success-team?
- Are you satisfied with the support?
- Would you like more or less contact with the company?

And lastly
If you were to investigate what parameters that is relevant to study in order to distinguish whether a customer is pleased and get the expected value out of the tool, or dissatisfied and in the risk zone of ending their subscription, how would you tackle that problem?
Appendix C: Web survey - Six short questions about Funnel

Introductive text
This questionnaire is part of my Master's thesis work and aims to get a better understanding of your perceptions and experiences with Funnel as well as how satisfied you are with the product. The questions take approximately one minute to answer.

Thank you for your participation!

//Robert Åman
Uppsala University

Survey
1. In general, how satisfied or dissatisfied are you with the product? (1=very dissatisfied to 5=very satisfied)
2. Have you ever contacted Funnel for support or general questions? (Yes/No)
   (If ‘Yes’, then move on to question 2a-c, if ‘No’ move on to question 3)
   a. How easy is it to get your issues or questions resolved? (1=Extremely hard to 5=Extremely easy)
   b. How responsive is Funnel's staff to your thoughts or questions? (1=Not at all responsive to 5=Extremely responsive)
   c. How committed is Funnel's staff when it comes to helping you? (1=Not at all committed to 5=Extremely committed)
3. How likely are you to still be a customer in three months? (1=Not at all likely to 5=Extremely likely)
   a. Could you please elaborate on the reason behind your previous answer? (Qualitative question with long text answer)
4. How likely are you to recommend Funnel to a friend or colleague? (1=Not at all likely to 10=Extremely likely)
5. If presented to the opportunity, would you be willing to participate in a case-study? (Examples of Funnel's current case studies: https://funnel.io/case-studies) (Yes/No/Maybe)
6. Finally, how would you describe Funnel for a friend or colleague? (Qualitative question with long text answer)
9.4 Appendix D: Distribution of quantitative survey answers and the re-grouping of variables

Bar chart illustrating the distribution of the Satisfaction-variable from the survey. The white boxes in the bars show the amount of cases as well as the percentage of the whole.

Bar chart illustrating the distribution of the re-grouped Satisfaction-variable. Responses having values 1-3 have been re-grouped into the variable “Less satisfied” while the responses 4-5 have been re-grouped into the variable “Satisfied”. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
Bar chart illustrating the distribution of the NPS-variable from the survey. The white boxes in the bars show the amount of cases as well as the percentage of the whole.

Bar chart illustrating the distribution of the re-grouped NPS-variable. Responses having values 1-6 have been re-grouped into the variable “Detractors”, responses 7-8 have been re-grouped into the variable “Passives” and responses 9-10 have been re-grouped into the variable “Promoters”. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
Bar chart illustrating the distribution of the Likelihood to stay-variable from the survey. The white boxes in the bars show the amount of cases as well as the percentage of the whole.

Bar chart illustrating the distribution of the re-grouped Likelihood to stay-variable. Responses having values 1-3 have been re-grouped into the variable “Fairly unlikely/Unsure”, responses having value 4 have been re-grouped into the variable “Fairly likely” and responses having value 5 have been re-grouped into the variable “Extremely likely”. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
The results from a crosstabulation with NPS as row variable and Satisfaction as column variable. Shows the difference in count vs. expected count. A Chi-square test showed statistical significance (p<0.05). See diagram below for illustration in bar chart.

<table>
<thead>
<tr>
<th>NPS</th>
<th>Detractors</th>
<th>Count</th>
<th>Expected Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>1,1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Passives</td>
<td>5</td>
<td>29</td>
<td>4,5</td>
</tr>
<tr>
<td>Promoters</td>
<td>0</td>
<td>41</td>
<td>5,4</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>72</td>
<td>11,0</td>
</tr>
</tbody>
</table>

The white boxes in the bars show the amount of cases as well as the percentage of the whole.

Bar chart illustrating the distribution of the Satisfaction-variable between the NPS-segments. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
The results from a crosstabulation with **Likelihood to stay** as row variable and **Satisfaction** as column variable. Shows the difference in count vs. expected count. A Chi-square test showed statistical significance (p<0.05). See diagram below for illustration in bar chart.

### Satisfaction vs. Likelihood to stay

<table>
<thead>
<tr>
<th>Likelihood to stay</th>
<th>Fairly unlikely/Unsure</th>
<th>Fairly likely</th>
<th>Extremely likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Expected Count</td>
<td>Count</td>
<td>Expected Count</td>
</tr>
<tr>
<td>Less satisfied</td>
<td>5</td>
<td>1.7</td>
<td>4</td>
<td>2.9</td>
</tr>
<tr>
<td>Satisfied</td>
<td>8</td>
<td>11.3</td>
<td>18</td>
<td>19.1</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>13.0</td>
<td>22</td>
<td>22.0</td>
</tr>
</tbody>
</table>

Bar chart illustrating the distribution of the **Satisfaction**-variable between the **Likelihood to stay**-segments. The white boxes in the bars show the amount of cases as well as the percentage of the whole.

---

71
Bar chart illustrating the distribution of responses between Funnel’s customer segments from the survey. The white boxes in the bars show the amount of cases as well as the percentage of the whole.
9.5 Appendix E: Qualitative answers from web survey and their clusters

How likely are you to still be a customer in 3 months?

All responses have been color-coded and grouped into clusters, depending on the central theme of each response. The clusters can be seen below, as well as what ratings that accompanied the responses in each cluster. All of the answers are also presented in their original form, grouped according to their rating.

Clusters

5  Continous development
5  Saves time
5  Superior to competitors

4, 5  Works as promised, good product and service, usefulness, need won't change
4, 5  Functionality
4, 5  Dependent
4, 5  Lock in

3, 4  Depends on other persons/factors
3, 4  Looking at other options
3, 4  Expect improvements

2  Problems with the product or set up

5s

- A good product, constantly improving, listen and work on feedback.
- Because we have already a yearly contract :)
- Bra verktyg och väldigt hjälpsam kundservice!
- Easy to use dashboards that make it really easy to get "glanceable" metrics and keep me in top of my paid metrics.
- Easy to use, quick to monitor result, the forecast function is great.
- Funnel allows me to add cost data across all channels, incl. offline funnels, which I cannot do to the same extent in other tools. Cost are reasonable too.
- Funnel has not been a static product and team, they are able to adapt and improve along with our business which is great.
- Good support and very responsive to questions.
- Good tool to keep track of our KPIs.
- Great customer service and a great product
- Great customer service. Easy implementation to your costs overview. (I think that there aren't many competitors as well..? which makes the service from Funnel unique.)
- Great features (which I need) and exceptional customer service
Great tool
I solves a big issue: implement easily cost data from different channels withing GA
I'm handling all my reporting with Funnel, it saves me tons of time on a weekly (maybe even daily) basis
It is just the right tool/platform for us with the best service!
It saves us 10 hours a week
It's a really helpful tool, really inexpensive and gives lots of insight.
It's the easiest and most efficient tool that allows me to import marketing costs in Google Analytics. All for a reasonable price.
My needs for a product like Funnel won't change anytime soon as I see it.
Provides the insight we are looking for
Replacing manual processes with automation
Saves time! Good service!
Service is very useful for us.
The fact that we are able to see our performance in different channels and markets in a quick way
The product realy suits in our business, and every problem we have had, was solved in less than 3 hours.
The solution is useful for us.
There aren't a lot of other options out there for this type of cost reporting
There is no better API to control costs vs GA Data actually available in the market that I know of.
This is the best dashboard tool we've used this far.
Time saving product with at an OK price level
We are highly satisfied with the product and the service.
We depend on the service
We find funnel to be a very useful tool for comparing key metrics from all our advertising campaigns. We also find the customer service to be quite outstanding and are very pleased with the fact that Funnel willingly ads new features to further help improving the tool after feedback from us as a customer.
We like Funnel due to the reason that it collects all data in one place. It's easy and convenient.
We like the program
We like the service - and it makes reporting easy.
We use Funnel on a daily basis. It's our reporting tool.

4s
Company uses Funnel as Ad spend tracker
Funnel is a lightweight tool, which is easy to set up. But, we are a huge organization with an internal business intelligence team. We may ultimately bring some of this functionality in house. Not likely within 3 months though.
Great tool - very useful in daily monitoring across a wide range of online marketing channels
I don't see this changing in the near future.

It depends on a number of people in the organization whether we keep funnel or not - all the people who would be using the reports. They still haven't seen any reports so I can't be 100% certain we'll be using the service in three months.

It is good, but there are some limitations in the product; which should be upgraded in the next releases.

Our company group uses Funnel for every information everyday basis

The tool is very useful to us

Very important to get the cost data in Google Analytics for Reporting

Very useful tool

We provided the dashboard some of our clients

We the API connector of funnel for reporting purposes, so we are not likely to switch anytime soon because of the investments we did (development of a custom excel plugin).

We're still waiting to see some improvements to the platform (targets, year-on-year comparison, etc.)

Works as promised, so no need to switch / change to alternative.

3s

Currently working on the integration of a different data visualization & integration platform

Depending on our needs and internal development

It depends on how the tool develops. By only linking impressions, clicks and cost to analytics transactions and transaction value does not give the full picture. What about attribution models etc?

It depends whether I still working at this company or not.

Not up to me.

Only being able to filter advertisers by Campaign Name is quite limiting

2s

Issues with data sync that weren't resolved - linkedin spend doesn't match in Google Analytics to what we are seeing when we login to LinkedIn -- didn't have time to keep troubleshooting and stopped relying on the data. Planning to cancel, as I don't have time to figure it out. LOVE the concept and wish it were more affordable so I could use it for all of my clients -- otherwise it makes more sense to manually upload the data to analytics each month - they don't have a need to daily syncing

We can't seem to get our account transferred over from our former agency to us directly. Plus, we have to upload LinkedIn data manually. And you don…

We didn't managed to set up all of our advertising accounts. It was at the beginning hard to get all access to them, then to do the mapping with campaigns in Funnel.
9.6 Appendix F: Test of linearity

Test of linearity indicating linear correlations or non-significant correlations between all variables except Recency and Created reports.