CUSTOMER SATISFACTION IN THE DIGITAL ERA

EXPLORING HOW DIGITAL SERVICE PROVIDERS CAN LEVERAGE PRE-EXISTING CUSTOMER DATA

ANNIE HARTMAN

JOHANNA ELMING

Bachelor Thesis

Stockholm School of Economics

2019

Customer satisfaction in the digital era

Exploring how digital service providers can leverage pre-existing customer data

Abstract:

This thesis reviews common methods and theories for measuring customer satisfaction for companies offering digitized services, with a focus on first movers. Specifically, it is investigated whether these theories are compatible with data-science models using pre-existing company data. To analyze this, a logistic regression analysis was made in order to assess the usability of customer satisfaction data, based on archive data received from a digital service in Sweden. The results led to a discussion regarding the disconnection between pre-existing data and conventional theories.

Keywords: Customer Satisfaction, Digital Services, Big Data, Logistic Regression,

Authors:

Annie Hartman (23965)
Johanna Elming (23639)

Tutor:

Patric Andersson, Associate Professor, Department of Marketing and Strategy

Examiner:

Hanna Berg, Research Fellow, Department of Marketing and Strategy

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics

Thank you to…

Patric Andersson for the engagement, support and valuable guidance.

Emelie Fröberg and Fredrik Blom for guidance and advice regarding *R*.

Friends and family for encouragement and support.

Contents

1. Introduction

During World War II, American economist and Nobel laureate Kenneth Arrow was responsible for producing long-range weather forecasts (Anderson, 2018). After a while, Arrow and his staticians found that their forecasts were no better than pulling predictions out of a hat. They then asked their superiors to be relieved of this duty. However, the Commanding General reported that he was well aware of this problem but needed something to base his planning on. This is a well-known example of uncertainty in predictions and the human unwillingness to accept it (Anderson, 2018).

In the digital era, where vast amounts of data are readily available, there are many advocators for its usefulness in marketing decision making, not the least when considering customer satisfaction. However, when examining common theories on customer satisfaction for digital services, a number of limitations can be observed. Given these implications, the aim of this thesis is to empirically analyze if pre-existing company data can be used to understand relationships between service quality and customer satisfaction for first mover digital service providers.

1.1. Background

1.1.1 The Digital era of Services

Electronic commerce has been reshaping the way of doing business in the last two decades (Daniel, Wilson, & Myers, 2002). It is defined as the buying and selling of information, products and services via computer networks (Kalakota & Whinston, 1997). This, has in recent years, resulted in a shift of the nature of services. By reaping off benefits from technological advances, changing customer behaviors and availability of data, digital startups are creating innovative and user-friendly alternatives to incumbents (D'Midio, Dorton, & Duncan, 2015).

1.1.2 The Swedish Digital Service Market

Sweden is one of the countries that have been leading in the startup movement during the last couple of years (GP. Bullhound, 2016). There has been a breakthrough with mobile applications offering services digitally, replacing traditional services and consultations. These digital alternatives are saving both time and money for consumers. *KRY* and *Min Doktor* are two examples of this phenomenon, both of which are offering healthcare consultations through online communication in the form of video calls with certified doctors. This new era of digital services has been proven successful among the Swedish customers. *Min Doktor* more than doubled their sales of digital healthcare to Swedish county councils between 2016 and 2017 and during the same period *KRY* managed to increase their number of health care consultations with 1600 percent (Lennen Merckx, 2018). The Swedish market has not only been penetrated by

innovations in digitalization within the health care sector. Numerous of other industries have also digitized their services, for example within banking and law. Often these companies are first movers leaving them with hundred percent of the market share within their line of work.

1.2. Problem Area and Research Gap

Digitalization does not only offer consumers better, more personalized and efficient services, it also provides businesses with large amounts of data about their customers. Most marketers are confident that big data will play a crucial role in achieving ambitious commercial targets (Almqvist, 2018). At the same time, the theories and models on how to track and measure business success are not fully adaptable to e-commerce (Xu & Quaddus, 2009). This may result in situations where digital service providers do not know how to efficiently collect and measure data in order to develop their businesses.

For digital services, one could argue that it is essential to follow the so-called *Golden Rule of Forecasting*. This rule advocates sophisticated prediction models that incorporate all available information assuming that having access to more data leads to better predictions (Armstrong, Green, & Graefe 2015). This is however somewhat paradoxical since there is also evidence saying that ignoring part of the information improves forecasting, and the fact that simpler models often do well (Almqvist, 2018; Wübben & Wangenheim, 2008). On this topic, an interview article with professor David Sumpter presents several examples from *Google* and *Facebook* revealing that artificial intelligence often make much worse decisions than humans (Byttner, 2019). This is further supported by the classic principle *Occam's razor*, also called *the law of economy*, stated by the scholastic philosopher William of Ockham in the 14th century. The principle states that out of two competing theories, the simpler explanation of an entity is to be preferred (Duignan, 2018). Having nuanced expectations on predictions from big data is of essence, in order not to forget about the importance of simple managerial heuristics[1] (Almqvist, 2018; Persson & Ryals, 2014; Wübben & Wangenheim, 2008). At the same time as data collection is being praised, it is difficult to assess the actual quality of the predictions it can enable.

Drawing on the possibilities of the large amounts of data accessible to digital businesses, one must consider the methods to measure success from these sources. Customer satisfaction as a result of service quality and expectations serve as basis for strategic decision making. It is crucial to understand customer satisfaction in a competitive market. Satisfied customers have a great effect on customer retention, something that has become even more important in the digital era due to low switching

[1]Businesses' rules of thumb for handling specific tasks in their environments (Persson & Ryals, 2014).

costs and more available information about alternatives. Customer satisfaction and service quality, in this thesis, is defined as:

> **Customer satisfaction:** The number of customers, or percentage of total customers, whose reported experience with a firm, its products, or its services (ratings) exceeds specified satisfaction goals. It is generally based on survey data and expressed as a rating (Farris, et al., 2015).
>
> **Service quality:** The degree and type of discrepancy between the perceptions and expectations of users suggesting that they all, in general, employed similar aspects of service by which quality could be assessed (Parasuraman, Zeithaml & Berry, 1985).

To understand what this means for digital businesses, following is a review of common customer satisfaction and service quality theories.

In 1977 the *Expectancy-Disconfirmation* model was introduced (Oliver, 1977), drawing on the *Adaptation level theory* from 1964 (Helson, 1964). These frameworks imply conscious comparisons between a cognitive state prior to an event and a subsequent cognitive state (Yüksel & Yüksel, 2008). In 1965, the *Dissonance Theory* was developed, which states that disconfirmation of expectations creates a state of psychological discomfort (Cardozo, 1965). One of the problems with these theories is that they presuppose that consumers have precise expectations drawing on experience prior to the service. Without these experiences dis/confirmation of expectations cannot occur. The theories in general become less meaningful in situations where consumers do not know what to expect until they experience the service (Yüksel & Yüksel, 2008).

In a review of customer satisfaction theories, Yüksel and Yüksel (2008) explain that health care services in particular are experiential by nature, and that they contain a high percentage of experience and credence properties. By adding the digital dimension, it is even more unlikely that the customer will have adequate prior experiences to contrast against (Yüksel & Yüksel, 2008).

Expectations of a product may not correspond to what is desired and valued in a product, and thus, values and attributes may be better comparative standards as opposed to expectations. (Yüksel & Yüksel, 2008) One of the most widely used models to assess attributes to customer satisfaction for services is *SERVQUAL* which was introduced by Parasuraman et al. in 1988 (Parasuraman, Zeithaml, & Berry, 1988). However, traditional service quality dimensions cannot directly be applied to digital services. Thereby, *e-SERVQUAL* was developed by identifying four dimensions; efficiency, reliability, fulfillment and privacy (Zeithaml, Parasuraman, & Malhotra, 2005). After that, *WEBQUAL* was introduced with twelve dimensions[2] (Lociacono, Watson, &

[2] See appendix 1 for dimensions.

Goodhue, 2000), *and .comQ* with four factors of website design, reliability, privacy and customer service (Wolfinbarger & Gilly, 2002). These theories are relevant for digital service providers.

Customer satisfaction as outcome data is often measured by asking the customer how satisfied he or she is. To best of the authors knowledge, there is no other common method to directly assess a customer's experienced satisfaction. Using questionnaires to assess service quality is also common. However, since there are so many dimensions of service quality it is a time consuming and costly method. Companies would benefit if they could use pre-existing data instead, something that digitalization and big data could make possible.

There are studies that analyze the use of archive data to assess company success. It consists mainly of two studies conducted on the Chinese market; one in 2015 by Hualong Yang, Xitong Guo and Tianshi Wu, and another one in 2018 by Yefei Yang, Xiaofei Zhang and Peter K.C. Lee. Both studies analyzed statistical data from the healthcare platform "Good Doctor Online", which was founded in 2006 and is one of the biggest online health communities in China (Yang et al., 2015). By examining different attributes of service quality from archive data they found that variations in service delivery behavior affected satisfaction. It would be of great value if this could be applied for digital service providers, as less resources would have to be invested in quality surveys.

However, there is little understanding of the factors that affect customer satisfaction within Internet-based services (Khalifa & Liu, 2003). As previously mentioned, there are many new entrants to the digital service market that are adding value by complementing or substituting traditional services by technological solutions that use less of physical resources. Digital services are superior to traditional incumbents by nature (Khalifa & Liu, 2003), and therefore it can be problematic for businesses to assess what actually makes their service better. Explicitly, for the case of digital health care, it is possible that the patient enters a state of satisfaction solely by booking the consultation. Meaning that they are satisfied, prior to the actual consultation, just by knowing that it is a more efficient alternative. Considering this, it would not be surprising if traditional theories of customer satisfaction do not match these business models. It is thereby evident that there is a research gap in consumer evaluation theories for newly started digital services (Parasuraman & Grewal, 2000).

1.3. Purpose and Research Question

Drawing on the background and problem discussion, the purpose of this essay is to empirically analyze how newly started digital service providers can leverage their pre-existing data to understand the relationship between service quality and customer

satisfaction. Based on this stated purpose, the following research question was formulated to be discussed from a marketing perspective:

1) How can digital companies measure the antecedents of customer satisfaction using their pre-existing data and available theories?

The question will be answered in regard to first mover digital service providers whose customers thereby have no prior experience to the service and where competition is nonexistent.

1.4. Delimitations

To make the analysis of this thesis manageable, delimitations had to be made. Geographically, this thesis is limited to analyze the Swedish market. To represent the first mover digital service providers one company was chosen.[3] The chosen company is a digital veterinary service that was launched in 2016. [4] According to their website they boast an impressive 98% customer satisfaction rate. Even though the company possesses large amounts of data about their service they have not adopted any certain theories or models on how to surveil relationships between service quality delivery and customer satisfaction, which seems to be a common issue among companies (Schwagger & Mayers, 2007). The data provided by the company is investigated in the light of common customer satisfaction and data prediction theories.

The company was selected for several reasons. Firstly, the company is believed to reflect a typical digital service company that is a pioneer in its field, which is the type of company that this thesis aims to analyze. Secondly, the company possess a big amount of data, including data on measured customer satisfaction and they agreed upon sharing the data for the purpose of this thesis. Lastly, the company is thought to reflect a first mover as they were the only available alternative on the Swedish market up until February 2019. The company is fast growing and have increased their number of users drastically since their launch. Arguments for selecting another company to represent first mover digital service providers are valid and further discussed in section 3.3.

The digital service offers customers the opportunity to obtain a consultation from a certified veterinarian through video calls, phone calls and direct messages between 07:00 and 24:00, every day of the week. The service works as a more efficient way of visiting the veterinarian physically and gives customers the opportunity to obtain a

[3] Problems with extrapolating one case on a bigger issue discussed in section 5.2.

[4] For confidentiality purposes and protection of company interests, the subject company is not mentioned by name in this thesis. Therefore, the sources regarding the company are excluded in the reference list. The authors assure that the information from primary sources at the company and secondary sources regarding the company are true and reliable. The primary sources consist of the provided data itself and interviews with the chief technology officer, a programmer and a veterinarian. The secondary source is the company website.

consultation without having to leave the comfort of their homes and dealing with the logistical and stressful difficulties associated with transporting their pet. Further, the company is cooperating with every insurance company in Sweden that offer animal insurance, allowing all pets with an insurance to be consulted without charge, resulting in pet owners saving both time and money.

Customers register their pets using BankID, a personal and easy service for secure electronic identification. After that, available times for consultations appear, which is usually approximately 15 minutes away. The customers book a time and register their case as well as information about their pet. When it is time for the consultation the customer is contacted by the veterinarian.

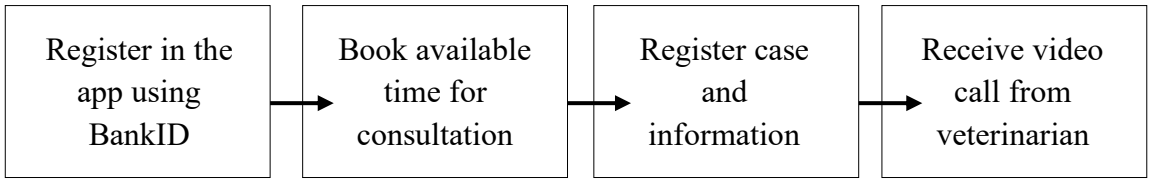

Figure 1. The digital veterinary service process

When discussing the high level of customer satisfaction with one of the veterinarians working at the company, her understanding was that some of the crucial factors contributing to the high results are that the service creates a feeling of exclusivity for the customer, that it is very easily accessible, the veterinarians are always happy to help, and the fact that customers are satisfied with the concept of the service itself. She believed that the rare cases of customers giving low rates are due to the fact that people sometimes expect their pet to be treated in a way that the veterinarian does not agree on, meaning that their expectations are not met when it comes to treatment.

To make a fair analysis for the purpose of this essay, it was chosen to delimit the data in the following way:

- Only video consultation: Video calls are the only consultations where the customers are asked to rate the service and is also the richest medium used in the service.
- Only dogs: A definite limitation to the primary data is that it only includes cases were customers have been consulted on dogs. This is to create a fair analysis and not compare different species, such as dogs and cats, considering that they must experience substantially different health problems. The choice of analyzing dogs is explained by the fact that dogs represent a majority of consultations. Further, dogs are the second most common pet in Sweden after cats (Agria Djurförsäkring, 2017b). Also, 9 out of 10 dogs have veterinary insurance while less than 5 out of 10 cats do (Agria Djurförsäkring, 2017a), making dogs a more interesting subject in this context.

1.5. Expected Contribution

The expected contribution of this thesis is mainly attributed to marketing, considering the fact that it will help digital companies with their customer satisfaction analysis. This is done by exploring if pre-existing data of service quality can lead to valuable conclusions about customer satisfaction. This, in turn, can provide companies with a more efficient and less costly method of measuring the antecedents of customer satisfaction and a more representative view of what customers value in their service.

1.6. Disposition

The remaining content of this Bachelor thesis is divided into five chapters. Previous research and theories related to the topic are presented in chapter 2. Following is chapter 3 which explains the methodology used in this study and discusses reliability and validity. Further, chapter 4 presents the findings from the data and an analysis of these. In chapter 5 the findings are discussed, and conclusions are stated together with a discussion of potential shortcomings and proposals for future research. The essay ends with references and thereafter an appendix.

2. Theoretical Framework

This chapter presents the literature on online services, service quality, customer satisfaction and implications of data that is examined to gain a better understanding of the chosen topic. First, models and theories of customer satisfaction are presented to understand service quality and its impacts. Secondly, literature on the implications of datafication is presented, in order to gain a better understanding of why data collection is important as well as understanding why it is not obvious that it should be considered superior to using traditional managerial heuristics. Lastly, previous research on digital services are reviewed.

2.1. Customer Satisfaction

To obtain competitive strengths it is crucial to measure service quality and customer satisfaction (Parasuraman, Zeithaml, & Berry, 1994). This is especially important for services within the health sector because patients experience high personal importance when using these services (Yüksel & Yüksel, 2008). This is also true for the firm examined in this essay even though it is within the animal health sector, since the health of one's pet is of high personal importance. There are several previous studies providing findings relating the service delivery process to customer satisfaction.

Parasuraman et al. (1985) found that services are intangible, heterogenous, and inseparable. Therefore, the intentions of the service provider can be significantly different from what the customer receives. Thus, the service delivery can affect the service quality, so when the customer evaluates the service, the satisfaction depends on the delivery process. Thereby, service quality is the result of a consumer's expectations compared to the actual service performance (Parasuraman, Zeithaml, & Berry, 1988).

Satisfaction judgment is related to the expectation-disconfirmation approach. Customers form expectations about a service before purchasing, which affects perceived quality. From those expectations, the customers then form a disconfirmation or satisfaction outcome (Anderson & Sullivan, 1993).

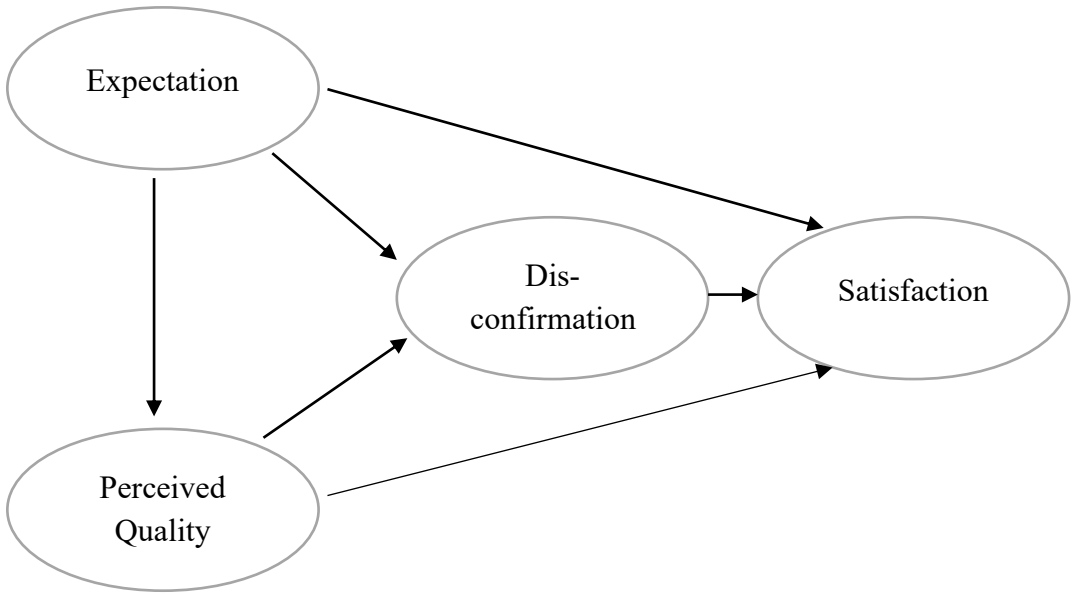

Figure 1. Satisfaction Formation

Oliver, referred to in Anderson & Sullivan (1993)

SERVQUAL was introduced in 1988 as a multi-item scale to measure the service quality process. The model has five dimensions; *Reliability*, *Responsiveness*, *Tangibles*, *Assurance* and *Empathy* (Parasuraman et al., 1988). Since traditional services are substantially different from online services in terms of fundamental quality dimensions, the above-mentioned theories are not adequate to measure service quality and customer satisfaction for online services. Previous research of internet-based services has been focused on the technology behind the services rather than looking at the delivery process and interaction between providers and customers (Yang, Guo, & Wu, 2015). Another aspect of perceived quality is that customers' characteristics fundamentally affect satisfaction. Differences in customer characteristics has moderating effects on satisfaction levels. These characteristics can be demographic referring to a customer's age, gender, and social and economic status (Anderson, Pearo, & Widener, 2008). They can also be situational meaning that they include purchase uncertainty, cost, purchase importance and length of customer relationship (Wangenheim, 2003). The length of a customer relationship also has significant impact on customer satisfaction. This, as trust development is a dynamic process meaning that it changes over time as the relationship between the service provider and the customer gradually develops through the interactions (Lewicki, Tomlinson, & Gillespie, 2006.; Lewicki, Mcallister, & Bies, 1998).

2.2. Big Data

The methods used for information search, payments, trade and shopping has changed during the past decade due to data-science algorithms and models. Along the way there have also been changes in how businesses use individual-level consumer data, and the way business transactions are created, documented, regulated, and analyzed (Giesecke, et al., 2018). Even though big data is a ubiquitous term today the definitions are broad and uncertain. The most common way of describing big data is by categorizing the aspects into a number of V:s (Gandomi, & Haider, 2015). Laney introduced *The three V:s* in 2001; *Volume, Variety* and *Velocity*. What a big volume is, is depending on the context. *Variety* is referring to the collection and types of data and the underlying sources, whereas *Velocity* measures data turnover of how quickly data is collected and being acted upon (Laney, 2001). Since Laney, there have been many additions from commercial actors, such as *International Business Machines*'s *Veracity* (IBM Big Data & Analytics Hub), *Statistical Analysis System*'s *Variability* and *Complexity* (SAS Insights, 2019) and *Oracle*'s *Value* (Oracle, 2013). The support of these V:s being essential in producing forecasts to achieving commercial objectives is manifold, forecasts extrapolate what is already known into the future and hope that the original relationships still hold (Almquist, 2018).

2.3. Critique Against Data-science Models

There are many data-science models that operate on information consisting solely of customers' past purchase behavior. In a study by Wübben and Wangenheim in 2008 two models of this kind, *Pareto/NBD* and the *BG/NBD* were analyzed and compared to simple managerial heuristics to examine the performance in predicting customer activity and purchase level. Logistic regression models provide promising approaches in determining these probabilities and the implementation of mentioned models has been recommended on a large scale. Both models are considered attractive because they (1) make forecasts of individual's future purchase levels and (2) operate on past transaction behavior. However, the authors find no clear evidence for the superiority of these models for managerially relevant decisions in customer management compared to simple managerial heuristics that are still commonly applied. The authors present reasons to believe that simple heuristics may work better than complex strategies for various types of tasks, even though they often require less information and computation. According to studies, experienced marketing managers are likely to be accurate in their intuitive skills because their environments are characterized by two properties that are essential for learning: repetition and feedback. A survey done in 2002 by Christian and Timbers reveals that 45% of corporate executives rely more on instinct than figures and facts in running their firm. Even though this survey was done 17 years ago, and the statistical models have high acceptance in the academic community, it is important to

shed light on how these methods can make a good fit in improving marketing decision making. Wübben and Wangenheim found that the focal stochastic models were limited, or as well performing, compared to the managerial heuristics in predicting customer behaviors. Their result was devastating for the key results of the NBD/Pareto model while the heuristics used by the examined firms worked well. This should be of concern for academics since these models have not found their way into managerial practice. It is important to know under which circumstances the predictions of these models can be trusted in producing accurate forecasts that outperform managerial heuristics. For example, the authors raise the question:

> Can any model, as sophisticated as it may be, make reliable forecasts for a customer who has conducted only one transaction with a supplier? (Wübben & Wangenheim, 2008)

To the best knowledge of Wübben and Wangenheim, no work has addressed the question of how many transactions a customer needs to have conducted before reliable forecasts can be made (Wübben & Wangenheim, 2008). This currently seems to hold true. As long as the issues regarding data-science prediction models exist, managerial heuristics cannot be ignored a priori. PhD Student Gustav Almqvist at Stockholm School of Economics describes that big data forecasts can be seen as susceptible to prediction uncertainty and complexity; consequently, big data alone does not guarantee better predictions. To ensure the value in data collection and improve marketing decision making based on that, there is a need for evaluating the quality of the prediction models as much as the data itself. Almquist concludes:

> There can be more to the interaction between a specific prediction model and a particular environment than what first meets the eye. (Almquist, 2018)

2.4. Previous Research on Customer Satisfaction within Digital Services

Research done on digital services in relation to customer satisfaction includes different topics, for example; the relationships between e-SERVQUAL and customer satisfaction in online stores (Lee & Lin, 2005), the relationships between service quality, customer satisfaction, customer trust, and loyalty to e-banks (Chu, Lee & Chao, 2012), and the key factors that determine customer satisfaction within governmental e-services (Alawneh, Al-Refai, & Batiha, 2013), to name a few.

However, to the best of the authors' knowledge, previous research on the topic of customer satisfaction in regard to online health consultations is limited. It consists mainly of the two studies conducted on the Chinese market by Yang et al., 2015 and Ynag et al., 2018. In these studies an online healthcare consultation was defined as the use of information and communication technologies by remote patients on the basis of

instant and non-instant technologies to communicate with doctors and acquire medical care via a website that generally include information on consultations, records, community forums and other information (Piete et al., 2014; Kamis et al., 2014). Customer satisfaction is measured differently in the two studies. However, in both, it refers to the patient's psychological state immediately after the doctor has dealt with their medical problem. When patients have completed their online consultation, they can give ratings about the physician's service quality immediately. In Yang et al., (2015), this is measured as a rating. In Yang et al., (2018), the fact that the doctor receives a virtual gift or not is used as a proxy for satisfaction. If a customer does not give the consultation a rating or gift, that is interpreted in the data analysis, for both studies, as the customer not being satisfied.[5] Customer satisfaction is thereby measured based on the customer's actual rating, which is the conventional method. However, the studies measure service quality based on transaction, behavioral and demographic data, which in many ways goes against the common theories of how service quality can be measured. This since these theories have mainly been developed to be used as basis for quality surveys.

2.4.1. Study by Hualong Yang, Xitong Guo & Tianshi Wu (2015)

The study conducted in 2015 investigated if response speed and interaction frequency affect customer satisfaction, as well as what the moderating effects of a patient's disease risk is on the relationship between the service delivery process and satisfaction. Data was collected from 16 disease categories with 150 doctors from each category, resulting in 2112 consultations from doctors (Yang et al., 2015).

To analyze the data and test the hypotheses, Yang et al. (2015) used an empirical model. The variables in the model are not normally distributed due to the large variance in independent and dependent variables. The model is therefore adjusted to a linear logarithmic model:

$$\begin{aligned} Ratio_i = a_0 &+ a_1 log(time_i) + a_2 log(frequency_i) + a_3 log(title_i) \\ &+ a_4 log(letter_i) + a_5 log(gift_i) + a_6 log(contribution_i) + a_7 risk_i \\ &+ a_8 risk_i * \log(time_i) + a_9 risk_i * \log(frequency_i) + u_i \end{aligned} \tag{1}$$

$a_1$to $a_9$are the parameters to be estimated, and $u_i$is the error term associated with observation i. The variables risk*log($time_i$)and risk*log($frequency_i$) are interaction terms[6] (Yang et al., 2015).

[5] Critique can be raised against the fact that Yang et al., (2015), and Yang et al., (2018), do not provide any empirical evidence that provides an explanation for this assumption.

[6] An interaction term shows the moderating effects of a situational variable on a relationship between an independent and a dependent variable (James et al., 2017).

Table 1 presents the descriptive statistics and correlations of the major variables and the results of the model estimated by *ordinary least squares*, OLS (Yang et al., 2015).

Table 1. **Results for Yang et al. 2015**

Hypothesis	Variable	Measurement	Results	Interpretation
H1a: The physicians speed hypothesis	$time_i$	A physician's response speed	$a_1 = -0.001$ $t = 2.260$ $p < 0.05$	Positive effect on customer satisfaction, H1a supported
H1b: The interaction frequency hypothesis	$frequency_i$	Interaction frequency between a physician and a patient	$a_2 = 0.049$ $t = 9.656$ $p < 0.01$	Positive effect on customer satisfaction, H1b supported
H2a: The response speed/customer satisfaction for high disease risk hypothesis	$risk_i * time_i$	Interaction term for response speed with regards to illness severity	$a_8 = -0.074$ $t = -4.480$ $p < 0.01$	Response speed more influential for high-risk diseases than for low-risk diseases, H2a supported
H2b: The interaction frequency/patient satisfaction for high disease risk hypothesis	$risk_i * frequency_i$	Interaction term for interaction frequency with regards to illness severity	$a_9 = 0.038$ $t = 3.432$ $p < 0.01$	Interaction frequency more influential for high-risk diseases than for low-risk diseases, H2b supported

2.4.2. Study by Yefei Yang, Yefei Yang, Xiaofei Zhang and Peter K.C. Lee (2018)

The dataset that was used for the second study covered two years of online healthcare consultations representing 77 248 patients and 93 doctors. The data was collected from two categories of patients, those suffering from lung cancer and those suffering from the common chronic disease diabetes, to represent different levels of illness severity (Yang et al., 2018).

The aim of the second study was to analyze if doctors' service delivery behaviors comprising a) service content, b) depth of interactions, and c) response time, during a first period consultation determines whether a medical case has a single consultation or multi-period consultations. The study also analyzed how the doctors' service delivery behaviors impacted patient satisfaction in different consultation periods. Relationships between companies and customers are built over time. Therefore, the relationships between antecedents of customer satisfaction and customer satisfaction rates are analyzed in single-period and multi-period consultations. Whether a customer is a first-

time user or returning customer therefore acts as a situational mediator of the effects on the dependent variable. Finally, the purpose of the study was also to examine how the severity of an illness moderates the associations between doctors' service delivery behaviors and patient satisfaction (Yang et al., 2018).

Further, a logistic regression of patients' decisions on subsequent-periods consultation was made. The regression was made with the dependent variable "patient's decisions on subsequent-periods consultation" and it was divided into two models. The first model included only the control variables and the second model included the control variables together with the independent variables. Finally, Yang et al (2018) conducted two hierarchy regressions to be able to test the moderating effect. The first hierarchy regression was for patient satisfaction in first period consultation and the second hierarchy regression was for patient satisfaction in subsequent-periods consultations. The results of the study are represented in *Table 2* below.

Table 2. **Results for Yang et al. 2018**

Hypothesis	**Variable**	**Measurement**	**Results**	**Interpretation**
H1: Response time positively affects patients' decisions as to whether they will continue onto subsequent periods of consultations	RT	Average time interval between patient questions and doctor answers in one consultation	$\beta_{RT} = -0.0006$ $p < 0.01$	Long response time can persuade patients not to continue onto subsequent periods of consultations
H2: The depth of interaction positively affects patients' decisions regarding whether they will continue onto subsequent periods of consultations	*DI*	The number of interactions in one consultation	$\beta_{DI} = 0.488$ $p < 0.10$	Deep interaction can persuade patients to continue onto subsequent periods of consultations
H3: Service content affects patients decisions regarding whether they will continue onto subsequent periods of consultations	*SC*	The word count of a doctor's response in one consultation	$\beta_{SC} = 0.0061$ $p < 0.01$	Detailed service content can persuade patients to continue onto subsequent periods of consultations
H4: Response time, depth of interaction, and service content significantly and positively influence customers satisfaction with first-period consultation	*SC* *DI* *RT*	See measurements for H1, H2 and H3	$\beta_{RT} = -0.00457$ $p < 0.01$ $\beta_{DI} = 0.0048$ $p < 0.01$ $\beta_{SC} = 0.000587$ $p < 0.01$	Response time has a negative impact on patient satisfaction, detailed service content and depth of interactions have a positive impact on patient satisfaction

H5: Service content significantly enhance patient's satisfaction during subsequent-periods consultations, whereas depth of interaction and response time do not	*SC* *DI* *RT*	See measurements for H1, H2 and H3	$\beta_{RT} = 0.000588$ $p > 0.1$ $\beta_{DI} = 0.000204$ $p > 0.1$ $\beta_{SC} = 0.000366$ $p < 0.01$	Detailed service content positively impacts customer satisfaction in subsequent-periods consultations, but response time and depth of interaction have no significance
H6: Severity of illness moderates the relationship between response time, depth of interactions, and service content with first period consultations	*IS*RT* *IS*SC* *IS*DI*	Interaction terms for response time, service content, and interaction depth with regards to illness severity	$\beta_{IS*RT} = 0.00551$ $p < 0.01$ $\beta_{IS*SC} = 0.000149$ $p > 0.1$ $\beta_{IS*DI} = -0.00289$ $p < 0.1$	Severity of illness positively influences the relationship between responsive time and that of depth of interaction and customer satisfaction. Illness severity relationship between service content and patient satisfaction is not significant
H7: Severity of illness moderates the influence of service content in subsequent-periods	*IS*SC*	Interaction term for service content with regards to illness severity	$\beta_{IS*SC} = -0.00068$ $p > 0.01$	Severity of illness in subsequent periods consultations negatively influences the relationship between service content and customer satisfaction.

3. Methodology

3.1. Research Approach

This thesis uses a semi-deductive, exploratory approach. This approach is suitable as the thesis aims to develop a better understanding, seek new insight and ask questions regarding a phenomenon in a new light (Saunders, Lewis, & Thornhill, 2003). There are some aspects of this thesis that are similar to a replication study since it aims to test if the unconventional methods used in Yang et al., 2015 and Yang et al., 2018 can be applicable in a related case. However, no hypotheses are presented in this thesis and the method used is a combination of two previous studies. These are reasons why this paper should not be seen as a replication study. When a study is based exclusively on quantitative research it can be difficult to determine whether or not it should be described as a case study. A case study has a focus on a bounded situation. One could argue that this thesis should be described as an instrumental case study since it focuses on using the case as a means of understanding a broader issue to allow generalizations to be challenged. However, the rules deciding whether a study is a case study or not are rather vague (Bryman & Bell, 2011). Other approaches could have been useful in investigating the research question.[7] The study design is aligned with the thesis purpose using an empirical study for understanding digital service quality, based on quantitative factors by using regression analysis. Regression analysis is suitable for equations involving large datasets to understand the underlying connections between service quality and customer satisfaction.

3.2. Model

Regression analysis requires careful model design. There are no general methods that instruct which covariates to include in a regression model. For the purpose of this thesis, a modified version of the models mentioned in the theoretical framework will be presented. The models presented in Yang et al., (2015), and Yang et al., (2018), serve as the basis of the model formed in this thesis due to the nature of the received data. Yang et al. 2015 uses a model estimated by OLS. This is perhaps the most common type of regression. This method was initially used in this thesis, but since the data was not normally distributed however, which is one of the underlying assumptions of OLS, the

[7] Other approaches could have been used. Most appropriate would probably have been to interview data analysts at digital service companies to assess how they use their data to measure customer satisfaction. However, that would not result in any empirical evidence of the data quality. Another option is to hold interviews with animal owners to assess the customer satisfaction and understand the implications of traditional satisfaction theories for digital services where customers lack experience. Lastly, another method could have been to do an experiment in a data collection software like Qualtrics. For example, by describing digital service scenarios and then evaluating the antecedents of customer satisfaction.

results were not significant. With this assumption not being fulfilled, one can use a non-parametric test instead (James, Witten, Hastie, & Tibshirani, 2017). Therefore it was chosen to dichotomize the dependent variable. This is further discussed in section 3.2.4. This thesis instead uses logistic regression[8], as in Yang et al. 2018, which is appropriate for binary outcome data. (Wübben & Wangenheim, 2008) Logistic regression relaxes the underlying assumptions of OLS. An S-curve is fitted to show the odds of a satisfied or non-satisfied customer depending on the covariates. (James et al., 2017) Rather than modeling Y directly, logistic regression models the probability that Y belongs to a particular category, in this study - satisfaction. For example, the probability of satisfaction given fast response speed can be written as;

$$\Pr(\text{Satisfied} = \text{Yes} \mid \text{Fast response speed}) \tag{2}$$

To avoid predicting values that are below 0 and above 1 the model uses the logistic function. Regardless the value of X, a sensible prediction will be obtained. The logistic function is stated as follows:

$$\frac{p(X)}{1 - p(X)} = e^{\beta_0 + \beta_1 X} \tag{3}$$

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 - e^{\beta_0 + \beta_1 X}} \tag{4}$$

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X \tag{5}$$

The expression on the left side in function 3 is called the odds and can take on any value between 0 and ∞, values close to 0 and ∞ indicate low and high probabilities of satisfaction, respectively. To estimate the unknown parameter coefficients, the maximum likelihood method is used. The mathematical details of this method are beyond the scope of this thesis but statistical software packages like *R* makes the process of fitting the model easy (James et al., 2017).

[8] Even though logistic regression in some cases can be seen as a more forgiving method compared to linear regression, it is not as widely used. The most likely reason to this is that the dependent variable includes the natural logarithm and that the procedure behind making the results comprehendable is quite unpleasant, it requires solid theoretical preparatory work. (Bjerling & Ohlsson, 2010)

There is no generally accepted measure for logistic regression that represents the R^2 in an OLS-regression.[9] Instead, a so-called pseudo R^2 can be used. However, it should be used with caution, since the variance in a dichotomous variable is depending on the distribution of the variable (Bjerling & Ohlsson, 2010). One common measure is McFadden's pseudo R^2 which is used in this thesis and defined as follows:

$$R^2_{McF} = 1 - ln(L_M)/ln(L_0) \tag{6}$$

where L_M is the likelihood for the model being estimated and L_0 is the likelihood function for a model with no predictors (Allison, 2013).

3.2.1. Dataset

The dataset used in this model was provided to the authors from its primary source, the digital veterinary service. The received archive data consisted of 13716 observations. The nature of the service that this dataset is derived from is explained in section 1.4 and goes well under the definition of an online health care service that was presented in section 2.4.

The data received only included consultations booked within one hour. This is since the customer has the option to choose between "next available time" and "pick a time" when booking a consultation. This option expresses two different types of consultations. A customer who chooses "pick a time" can book a meeting two hours away, or even two days away, this implies that he or she will not have to wait, but prefer to pick a time that is suitable.[10] To be able to measure waiting time, this thesis only examines consultations where the customer has waited less than one hour as this is thought to exclude the customers who chose "pick a time" and therefore were not affected by the waiting time. This decision was made in consultation with company officials. The observations of the received data cover a time span of five months from October 1st, 2018 until March 1st, 2019. The subject company started to collect diagnostics in a new way in the end of September 2018, which may or may not have effects on the data. To avoid introducing any bias to the analysis in this thesis, only data after September 2018 was employed.

Further, some of the data points had to be excluded for the modified model to work analogically with previous research. The initial sample size was reduced to 5826 observations in two steps. Firstly, after consulting with one of the veterinarians at the

[9] Based on the proportion of total variation of outcomes in a model, R^2 measures how well observed outcomes are replicated by a model.

[10] The authors did not receive data on the number of customers who chose the option "pick a time", it can be questionable whether this exclusion introduces bias or not to the data. However, from a statistical perspective, including observations where the customer scheduled the consultation on a different date would not be comparable with a customer who has waited five to sixty minutes.

digital veterinarian service it was decided to limit the data to illness cases of *Intestinal* and *Skin, Fur and Ears*. As dog owners have a hard time determining the severity of the illness of their dog themselves, it is more relevant to look at acute versus non-acute illnesses instead of illness severity as in previous studies. *Intestinal* is considered to be most acute since it can include symptoms like fever, uncontrollable urination and defecation and reduced general condition. Henceforth, these symptoms cause a lot of stress and worry for dog owners. The veterinarian explained how some customers are up all night caring for their dog with intestinal problems, just waiting for the service to open. *Skin, Fur and Ears* is thought to be one of the least acute problems. This is as it is often external issues that are more related to appearance and not the immediate health of the animal. With these two case types being the two major ones out of the 12 most common, this limitation more than halved the data size, but still kept it at a large amount of 5827 observations.

Secondly, extreme values were detected in the data set and have therefore been excluded. The outliers consisted of three observations which differed significantly from the other observations in the category "animal age", with two dogs registered as 498 years old and one as 43 years old, which are not plausible ages for a dog.

3.2.2. Demographics

In order to gain an understanding of the customers using the online veterinary service, data was analyzed from a demographic perspective. The data was analyzed according to five different distinctions; demographics for consultations for all users, consultations for those who have rated the service, for those who have rated the service below 4, those who have not rated the service at all as well as those consultations within the illness categories "Skin, fur and ears" and "Intestinal".

The data that represented the age of the customers as well as for the dogs was categorized in order to make it easier to interpret clear segments. The ages of the dog-owners were divided into six intervals.[11] After segmenting the customers, the patients (dogs) were categorized into five segments according to age groups[12] (Agria Djurförsäkring, 2010). The dogs were also categorized according to the illness their owner registered when booking the consultation. The categories available are the following: *Other*, *Eyes*, *Toxic*, *Urinal*, *Respiratory*, *Intestinal*, *Genital*, *Feeding & General care*, *Lameness*, *Skin, Fur & Ears*, *Wounds* and *Oral & Dental*.

[11] "Children" (ages below 18), "young adults" (between 18 and 30),"adult" (between 31 and 45), "middle-aged" (between 46 and 64) and "seniors" (65 years or older).

[12] "Puppy" (younger than one years old), "teenager" (1-2 years), "adult" (3-5 years), "middle aged" (6-9 years) and "senior" (10 years or older).

When analyzing the statistics for the demographic variables in the different segments, there were no large differences. For the sake of brevity, *Table 3* only includes the descriptive demographics for the consultation types *Intestinal* and *Skin, Fur & Ears*. For more detailed description of all data, see *Appendix 2.*

Table 3: **Demographics for *Intestinal* and *Skin, Fur & Ears* consultations**

Variable	Category	Descriptive
Panel 1: Gender of dog owner	Male Female	76% 24%
Panel 2: Age of dog owner	Max. Min.	86 years old 14 years old
Panel 3: Age of dog	Max. Min.	18 years old 0 years old
Panel 4: Illness severity	Intestinal Skin, fur & ears	51.42 % 48.58 %
Panel 5: Insurance	Free consultations Paid consultation	99% 1%
Panel 6: Type of customer	First consultations Not first consultations	89% 11%
Panel 7: Hours	Regular hours Uncomfortable hours	40% 60%
Panel 8: Waiting time	Long waiting time Medium waiting time Short waiting time	6% 26% 68%
Panel 9: Duration	Long duration Medium duration Short duration	75% 22% 3%

3.2.3. Data Preparation

In order to use regression analysis some data often needs to be formatted in a specific manner. All covariates in this thesis, except customer and animal age, were not in a desired format and therefore had to be manipulated in Excel. In the case of a qualitative covariate a common method is to use dummy variables[13]. Some of the variables have three levels, *response speed* and *consultation duration*. There will always be one less dummy variable than the number of levels (James et al., 2017). For example in *response speed*, the level with no variable is *slow response speed* and this level is known as the baseline, see below.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i = \begin{cases} \beta_0 + \beta_1 + \epsilon_i \;\; if\ ith\ consultation\ has\ fast\ speed \\ \beta_0 + \beta_2 + \epsilon_i \;\; if\ ith\ consultation\ has\ medium\ speed \\ \beta_0 + \epsilon_i \;\; if\ ith\ consultation\ has\ slow\ speed \end{cases} \tag{7}$$

3.2.4. Dependent Variable

The dependent variable is customer satisfaction. This variable is measured in the same way as in the previous studies by Yang et al., (2015). Immediately after completing a video consultation using the service, customers are asked to rate the consultation quality in the app and can grade the service on a scale of 1 to 5. The rate does not have any anchors, the customer can choose between 1, 2, 3, 4 or 5 stars. Using 1-5 is one of the most common methods to measure customer satisfaction, and can also be referred to as a Likert scale. A typical five level Likert scale uses the following format; 1. Strongly disagree; 2. Disagree; 3. Neither agree nor disagree; 4. Agree; 5. Strongly agree (Armstrong, 1987). By using one of the most used psychometric scales it was chosen to translate the customer satisfaction scale in this thesis to; 1. Strongly Dissatisfied; 2. Dissatisfied; 3. Neutral; 4. satisfied; 5. Strongly satisfied. The question asked was as follows:

> How satisfied are you with the consultation?

When plotting the data, it is discovered that 98% of the customers who have given a rating are satisfied or strongly satisfied, i.e. they have given a rate of 4 or more. 92% out of the customers who have rated the service have given a rate of 5, they are strongly satsified. Due to the fact that the data was not normally distributed it was chosen to dichotomize the dependent variable data, into dissatisfied and satisfied. In analogy with Yang et al., 2018, it was also decided to include the observations that lacked a score and

[13] A dummy variable is expressed as 1 or 0, it is binary. An example of a binary variable is gender, i.e. female(1) or male(0), (James et al., 2017).

treat these as dissatisfied. Considering that a 4 or 5 on the Likert scale can be treated as satisfied or strongly satisfied, it was chosen that all scores below 4 are considered to not be satisfied. Hence, being neutral is not considered being a satisfied customer. A dummy variable is used to measure satisfaction, and satisfied customers and dissatisfied customers are expressed as 1 and 0, respectively. In the regression the dependent variable is denoted Satisfied_customer.

3.2.5. Covariates

The covariates used in this study is listed below, along with explanations of their role in the model, their unit, and whether they are dummy variables or not. In the two models studied in previous research, *interaction depth* is an independent variable, measured as the frequency of doctor-patient interactions per case. This measure is not applicable for this study, as consultations are done through video and not chat or forum as in the other studies, instead the customer gets a response directly as the call continues. It is rare or even non-existent that one case for the online veterinary service includes several consultations. The study by Yang et al., (2018), introduces the independent variable *service content* which was also introduced in this model but modified as consultation duration.

- Response speed: This variable refers to the amount of time between a customer booking a consultation and actually receiving the video call from the veterinarian. This is an important reflection of the service's performance in efficiency. As this data was given to the authors in intervals labeled "short_wait", "medium_wait" and "long_wait" they were decoded into two dummy variables for "Response_speed_fast" and "Response_speed_medium" with slow response speed as a baseline.
  - Response_speed_fast: A dummy variable for fast response speed. Measurement: 0: Not fast, 1: The response speed was fast.
  - Response_speed_medium: A dummy variable for medium response speed. Measurement: 0: Not medium, 1: The response speed was medium.
  - When both dummy variables are 0, the response speed was slow.
- Service Content: This variable refers to the extent of the consultation, including offering the customer information on conditions, explaining the care and recovery process, and in some cases giving the customer a prescription or remittance for their animal. In previous studies word count of responses has served as a proxy for the quality on the service content. Thus, length of consultation is used as a measurement of service quality. As this data was given to the authors in intervals labeled "short_length", "medium_duration" and "long_duration" it was decoded into two dummy variables for "Consultation_duration_short" and "Consultation_duration_medium" with long consultation duration as a baseline.

- Consultation_duration_short: A dummy variable for short consultation duration.
  Measurement: 0: Not short, 1: The consultation was short.
- Consultation_duration_medium: A dummy variable for medium consultation duration.
  Measurement: 0: Not medium, 1: The consultation duration was medium.
- When both dummy variables are 0, the consultation duration was long.

3.2.6. Interaction terms

An important situational variable for the digital veterinary service was thought to be if the case type could be considered acute or non-acute. These situational characteristics are thought to influence the customer's emotional involvement with their animal's health condition and the potential threat that comes with it (Wangenheim, 2003). Analogically with Yang et al., (2015), and Yang et al. (2018), this suggests that the independent variables may be moderated to impact customer satisfaction more or less in these situations. A customer facing an acute problem may put more value into short waiting time than a long duration of the video call. A dummy variable was used for the acute level by decoding the case type "Intestinal" as acute and the case type "Skin, fur and ears" as non-acute, see section 3.2.1. for explanation. Acute cases and Non-acute cases are expressed as 1 and 0, respectively. The interaction terms were computed as follows.

- Acute_illness: A dummy variable for acute level.
  Measurement: 0: Non-Acute illness 1: Acute illness
- Consultation_duration_short*Acute_illness: An interaction term for the interaction between the variables Consultation_duration_short and Acute_illness.
- Consultation_duration_medium*Acute_illness: An interaction term for the interaction between the variables C_d_medium and Acute_illness.
- Response_speed_short*Acute_illness: An interaction term for the interaction between the variables Response_speed_short and Acute_illness.
- Response_speed_medium*Acute_illness: An interaction term for the interaction between the variables Response_speed_medium and Acute_illness.

3.2.7. Control Variables

In previous studies, the doctor's attributes and past online experiences were adopted as control variables. Since customers cannot make the choice of which veterinarian that will consult them, and due to the lack of data, it is not adopted for this study. The control variables used are instead demographic variables such as *customer age*, *customer gender*, *animal insurance* and *animal age*. Another control variable is *regular hours*, measuring if the consultation was done during regular working hours, 08:00-18:00 on weekdays, or not. The control variables, except the age variables, were coded as dummy variables.

- Regular_hours: If the consultation was done between 08:00 and 18:00 on a weekday or not.
  Measurement: 0: Not during regular hours, 1: During regular hours
- Customer_gender: If the customer is a male or a female.
  Measurement: 0: Male, 1: Female
- Customer_age: The age of the customer.
  Measurement: Σ Age
- Animal_insurance: If the customer has insurance for their animal or not.
  Measurement: 0: No insurance, 1: Insurance
- Animal_age: The age of the animal.
  Measurement: Σ Age

3.2.8. Variable Summary

Table 4. **Variable measurement (N=5826)**

Variable		**Description**	**Measurements**	**Statistics**
Dependent variable:	Patient satisfaction	The rating given by the customer, describing their degree of satisfaction	0: dissatisfied 1: satisfied	41.66% 58.34%
Independent variable:	Response speed	The average time interval between booking a time and a veterinarian calling, described as fast or medium	0: not fast speed 1: fast speed	68% 32%
			0: not medium speed 1: medium speed	74% 26%
	Consultation duration	The duration of a video call, described as short or medium	0: not short duration 1: short duration	97% 3%
			0: not medium duration 1: medium duration	88% 22%
Moderator variable:	Acute illness	Two types of illnesses, acute and non-acute	0: non-acute illness 1: acute illness	48.58% 51.42%
Control variable:	Regular hours	The time and day of the week that the consultation is made	0: not during regular hours 1: regular hours	60% 40%
	Customer gender	The gender of the customer	0: male 1: female	76%

			24%
Customer age	The age of a customer	Σ Age of customer	Min: 14 Max: 86
Animal insurance	Whether the dog has an insurance or not	0: not insured animal 1: insured animal	1% 99%
Animal age	The age of the dog	Σ Age of dog	Min: 0 Max: 18

3.2.9. Model Estimation

To investigate the relationships between the independent variables and customer satisfaction the following empirical model was created.

$$log(\frac{p(X)}{1-p(X)}) = a_0 + a_1(Regular\_hours) + a_2(Customer\_gender) + a_3(Customer\_age) + a_4(Animal\_insurance) + a_5(Animal\_age) + a_6(Acute\_illness) + a_7(Acute\_illness) * (Consultation\_duration\_short) + a_8(Acute\_illness) * (Consultation\_duration\_medium) + a_9(Acute\_illness) * (Response\_speed\_short) + a_{10}(Acute\_illness) * (Response\_speed\_medium) \quad (8)$$

$$p(X) = \frac{e^{a_0 + a_1(Regular\_hours) + \ldots + a_{10}(Acute\_illness) * (Response\_speed\_medium)}}{1 + e^{a_0 + a_1(Regular\_hours) + \ldots + a_{10}(Acute\_illness) * (Response\_speed\_medium)}} \quad (9)$$

Where a_1 to a_{10} are the parameters to be estimated and p(X) is the satisfaction probability.

3.2.10. Procedure

The computations were done by using the statistical computing software *R*.

In order to better assess the accuracy of the model using logistic regression it was first fitted using only part of the data set, to later examine how well it predicted the held-out data (James et al., 2017). Therefore, the data set was divided into two samples. 1500 randomly picked observations were excluded from the training set used in the regression. Hence the regression was made using 4326 observations. Thereafter a prediction was made on the remaining 1500 observations using the coefficients from the

regression to assess how well the model can predict the probability of customer satisfaction.

A commonly used method for fitting the data to a logistic model is to use the command g*lm*, a built-in command in *R*. The input is the dependent variable as a factor, the independent variables as factors, the control variables and the interaction terms. In *R* the dependent variable and the dependent variables for response speed and consultation duration were treated as factors using the function *as.factor*. In order to tell *R* to run a logistic regression rather than some other type of generalized linear model the argument *family=binomial* is passed in (James et al., 2017).

To view the results of the g*lm*-command the function *summary* was used. It yields an output table containing the estimated values of the coefficient representing each covariate, the standard error for each covariate, the z-values for each covariate and the p-values for each covariate. The z-statistic associated with each coefficient plays the same role as the t-statistic in linear regression. A large value indicates evidence against the null hypothesis. The table also yields an *Akaike information criterion*, AIC. This criterion is defined for models fit by maximum likelihood, it is a measurement of goodness of fit, it is useful when comparing models as a lower value indicates higher quality (James et al., 2017).

McFadden's pseudo R^2 is calculated in *R* by fitting the model and the null model which only contains an intercept. Then, the measure is calculated as follows by using the fitted model log likelihood values; *1-logLik(mod)/logLik(nullmod)* (Bartlett, 2014).

The model is presented hierarchically, first showing a model only including control variables, then introducing independent variables and lastly introducing the interaction terms. Thereafter, a confidence interval of 95% was used to determine the relevance of the variables. The function used in *R* for this is called *confint.* and yields a lower and upper limit of a 95% confidence interval.

To carry out the prediction on the held out data the built-in function in *R, predict(),* was used. The inputs were the fitted model and the data set of the held out 1500 observations. The type *"response"* was used to collect the predictions in a vector. Then, the accuracy rate of predictions was calculated. In order to evaluate the predictive power of the model, the error rate was compared against a naïve approach using a random output vector with an expected value equal to that of the training set of 58.34%. This test was performed 100.000 times.

3.2.11. Ethical Considerations

Ensuring personal integrity and protecting citizens' interests as well as the interests of the subject company are critical challenges when handling large sets of customer data (Helbing et al., 2017). In this essay, these matters have been dealt with respect by not

including any personal information, nor explicitly communicating any company information in exact numbers. As a wish from the subject company to protect them from communicating competitor sensitive data, all data for *response speed* and *consultation duration* have been coded in spans of "short", "medium" and "long". The authors have after the conduction of this study handed back the data to the subject company and the authors no longer have the data in their possession.

3.3. Reliability and Validity

The most common tools for measuring the quality of a study are reliability and validity. Reliability refers to the consistency of a model and is most commonly measured by using one of three methods; test-retest reliability, internal consistency or inter-rater reliability (Bryman & Bell, 2011). In this thesis, the internal consistency was used as a measure for reliability. More specifically, the dataset was split into two samples. 1500 observations were excluded in the first regression. Thereafter the regression was run again on held out data to determine prediction quality. This is a suitable method to assess model accuracy (James et al., 2017). Further, the reliability of the model is also supported by the fact that only p-values below 0.01 were accepted, rather than 0.05. Some of the false positive results reported in the academic community are attributed to the norm of treating p-values below 0,05 as statistically significant. High statistical power is further ensured by using a large amount of data (Dreber & Johannesson, 2018). These concerns are especially of interest when replicating studies. It might be argued that this study is a conceptual replication of Yang et al., (2015) and Yang et al., (2018), however it is not since it is not testing for the same hypotheses, simply investigating the same field and using these studies as a basis (Dreber & Johannesson, 2018). Even though this thesis is not a replication it is still important to consider the concerns of such a study. This, since some aspects of this thesis are similar to a replication study.

Validity refers to the extent of which the scores from a measure represents the variable intended (Bryman & Bell, 2011). The large amount of data, 5826 observations, used in the analysis increases the validity and trustworthiness of the model and also indicates a reasonable representation of the population of total customers. The data also represents consultations given over a longer period, five months, which further increases the validity of the model.

Considerations do however need to be made in regard to the fact that the data only represents customers who have had consultations using the video call function and does not capture the customers who have declined using the service. As the Hungarian mathematician Abraham Wald proved when analyzing bullet holes in aircrafts during the World War II; the bullet holes in the aircrafts that came back after a battle were not the most valuable for understanding how the planes should be armored. However, the planes that did not come back, and that were therefore not part of the observations were

of more value. This, since they could have indicated where the planes were most vulnerable (Denrell, 2005). This theory could be applied as critique against the validity of this model.

One might also argue that another company should have been chosen. The chosen company is within the health sector just like the companies in the studies by Yang et al., (2015) and Yang et al. (2018). However, it is questionable whether a model built on human healthcare can be applied to animal health care. There are also risks with handling archive data. The data was not collected with the intention of being used in this model which can be problematic when used in this format. Even though previously mentioned outliers have been excluded, there needs to be trust put into the data set that there is no bug or human error that has been missed in the observations.

4. Results and Analysis

4.1. Results from Regression

Table 5 shows the results of the model estimated by maximum likelihood. It presents the outputs of the regression run on the thesis model on a sample of 4326 observations.

The dummy variables for short and medium consultation time have negative coefficients. Short consultation time has a p-value below 0,001 and medium consultation time has a p-value below 0,001 only in the second model with a p-value below 0,5 in the third model. It is clear from *Table 6* that the covariates do not contain a zero in the confidence interval of 95%.

The dummy variables for fast and medium response speed have positive coefficients and are statistically significant only in the third hierarchical model. Fast response speed has a p-value below 0,01 and medium response speed has a p-value below 0,05. *Table 6* shows that the covariates contain a zero in the confidence interval of 95% in the second model.

Two variables out of the control variables are statistically significant; customer age with a negative coefficient and animal insurance with a positive coefficient. Customer age has a p-value below 0,001 in all models, but animal insurance only has a significant p-value in the second model below 0,05. However, *Table 6* illustrates that the covariate animal insurance contains a zero in the confidence interval of 95% in the second model.

Two interaction terms are also statistically significant; the interactions between fast response time and acute illnesses and the interactions between medium response time and acute illnesses. The coefficients are negative with a p-value below 0,05. *Table 6* illustrates that the covariates do not contain a zero in the confidence interval of 95%.

The prediction in R using the fitted model on the held-out data resulted in an overall accuracy of prediction rate of 58.33%. This means that there is an error rate of 41.67%. The test using a naïve approach had an average error rate of 48.94%.

Table 5. **Parameter estimates**
(N=4326)

Variable		**Mean (1)**	**Mean (2)**	**Mean (3)**
Control variables	*Intercept*	0.760***(4.065)	0.800***(3.525)	0.588* (2.189)
	Regular_hours	-0.017(-0.270)	-0.055(-0.852)	-0.060(-0.934)
	Customer_gender	0.117(1.578)	0.155(1.537)	0.111(1.480)
	Customer_age	-0.020***(-8.531)	-0.021***(-8.925)	-0.020***(-8.768)
	Animal_insurance	0.324.(1.941)	0.313.(1.861)	0.300.(1.779)
	Animal_age	0.002(0.268)	0.001(0.185)	-0.003(-0.334)
Independent variables	*Consultation_duration_ short*		-1.146***(-6.130)	-1.173***(-4.809)
	Consultation_duration_ medium		-0.360***(-4.800)	-0.248*(-2.311)
	Response_speed_fast		0.190(1.384)	0.523**(2.624)
	Response_speed_ medium		0.083(0.571)	0.474*(2.240)
Interaction terms	*Acute_illness*			0.433(1.322)
	*Consultation_duration_ short*Acute_illness*			0.001(0.003)
	*Consultation_duration_ medium*Acute_illness*			-0.239(-1.588)
	*Response_speed_short *Acute_illness*			-0.0633*(-2.297)
	*Response_speed_medium *Acute_illness*			-0.730*(-2.503)
Other	AIC	5807.8	5754.7	5741
	Pseudo R^2	0.01364828	0.2771277	0.02808512

Z statistics in parentheses, Significance code: “***”p<0.001, “**”p<0.01, “*”p<0.05, “.”p<0.1, “”p<1

Table 6. **Upper and lower limits of coefficients in a confidence interval of 95%**
(N=4326)

Variable	Lower (1)	Upper (1)	Lower (2)	Upper (2)	Lower (3)	Upper (3)
Intercept	0.394	1.127	0.356	1.247	0.061	1.115
Regular_hours	-0.142	0.107	-0.181	0.71	-0.187	0.066
Customer_gender	-0.28	0.262	-0.031	0.261	-0.035	0.258
Customer_age	-0.024	-0.015	-0.025	-0.016	-0.025	-0.016
Animal_insurance	-0.004	0.651	-0.018	0.642	0.030	0.606
Animal_age	-0.015	0.020	-0.016	0.019	-0.032	0.631
Consultation_duration_ short			-1.519	-0.785	-1.662	-0.702
Consultation_duration_ medium			-0.507	-0.213	-0.457	-0.037
Response_speed_fast			-0.080	0.458	0.132	0.916
Response_speed_ medium			-0.202	0.366	0.059	0.890
Acute_illness					-0.091	0.961
Consultation_duration_ short*Acute_ilness					-0.755	0.742
Consultation_duration_ medium*Acute_illness					-0.534	0.056
Response_speed_short* Acute_illness					-1.175	-0.094
Response_speed_ medium*Acute_illness					-1.304	-0.160

4.2. Model Analysis

When analyzing the results, several shortcomings were detected. The AIC was somewhat reduced when adding more variables in the hierarchy regressions. McFadden's pseudo R^2 is noticeably low in all models, which does not speak in favor of the model. Since the dummy variable for medium consultation time is statistically significant with a p-value below 0.01 only in the second model, the covariate for *consultation duration* is not determined. The dummy variables for *fast* and *medium response speed* for the covariate *response speed* are only statistically significant in the third model. Furthermore, the variables both contain a zero in the confidence interval in the second model. *Animal insurance* and the interaction terms for acute illness and response speed are determined by a p-value below 0.05, this does not act in favor for these variables as the literature states that 0.01 is considered a more acceptable level. Hence, all variables, except the control variable *customer age* can be considered to not contribute with any explanation value of customer satisfaction in this model. When using the fitted model to predict outcomes of the held-out data there was an error rate of 41.67%. Comparing this to the average error rate from the test of 48.94%, the model is 7.27% better on predicting the outcomes than the naïve approach.

The results from the regressions show some similarities to the results in previous studies. The results from the study by Yang et al. (2015) found that the coefficient for response time was negative, which was also the case for response speed in this study. Further, the results showed that the coefficient for response time in regard to illness severity was also negative, which is the same for response speed in regard to acute illness for this study.

As mentioned, there was one variable that was significant in the conducted models though; the control variable *customer-age*. The variable had a negative impact on customer satisfaction, meaning that the higher age a customer had, the less satisfied they were with the service. This aspect is interesting to analyze, since it does not necessarily have to do with the service provided but could perhaps be a question of attitude towards digitalization in general. From a company perspective this could either mean that they should consider targeting a younger segment, or that they should make the service more user friendly for an older generation that lack experience in the digitalized era.

Even though there are some indications of similar results in this study as the previous ones, there is a lack of statistical significance and prediction quality. Most likely, this is due to problems in the data set or the model itself. Considering that the model in this thesis was based on previous research within a relatively narrow field it would be difficult to argue that it is not based on relevant theories. Due to the nature of the data, perhaps an even more suitable model was created. Since customer satisfaction is measured commonly as a categorical variable, logistic regression can be seen as an improvement of the available model in Yang et al. 2015 that uses linear regression.

Even though several measures were taken to improve the model, the results presented few explanation values. Therefore, it is more likely that the problem is associated with the data itself, since it is archive data, and the theories used are usually applied to survey data. Additional analyses were performed, including the initial OLS, which, for the sake of brevity, cannot be displayed here.

5. Discussion and Conclusion

5.1. Discussion

Given the lack of research in the field of customer satisfaction theories for digital services as well as the evident critique against data-science models, this thesis tries to shed light on this open question. To understand how newly started digital service providers can leverage their pre-existing data to understand the relationship between service quality and customer satisfaction, common theories and previous study models were used on a seemingly big data set. This resulted in the following insights:

- A logistic regression model predicting customer satisfaction using data from a first mover digital service could not provide any significant explanation values.
- Since the most common theories within customer satisfaction and service quality were developed to be used in surveys and questionnaires there seems to be a problem with applying them on pre-existing company data.
- There is a demand for theory that unites customer satisfaction, service quality and data-science models with a high statistical power to offer value for marketing decision making, in order for companies to leverage their pre-existing data.
- Measuring customer satisfaction as a categorical factor is problematic in data-science models due to interpretation difficulties.
- The lack of explanation value from the model can be due to the fact that customer satisfaction requires prior experience, which the customer base of the digital service does not have, depicting a first mover disadvantage. There is a risk that first movers are resting assure that everything is going well and therefore not guarding themselves towards future competition.

5.1.1. The Disconnection between Pre-Existing Data and Conventional Theory

The studies conducted by Yang et al. in 2015 and in 2018, showed results that gave valuable information for Good Doctor Online. The results indicated which factors of a consultation that lead to satisfied customers. When analyzing the same parameters in relation to customer satisfaction for the Swedish online veterinary service, the results could not be used to determine what factors of the consultations were generating the high levels of customer satisfaction. A possible explanation to this is that the method used is based on theories that were developed for survey data. When a company collects this type of data, the questions are asked in a direct way providing answers that are pinpointing exactly what is investigated. When using archive data, one must, as in this thesis, adjust the data and the way it is interpreted by using proxies and assumptions.

This thesis uses common theories of customer satisfaction as well as metrics and methods recommended in the statistical literature. The model was drawn from two previous, relevant models, and improved with support from statistical experts. Despite this, no valuable conclusions could be drawn from the data set. This is not to say that the data collected is worthless. The data can be used for many other aspects, for example to see trends in diseases connected to time of year, or to track what days or hours are most common for customers to call. The data does however seem to be rather worthless when it comes to tracing what causes the customers to be satisfied and what they value in a digital consultation. Just like Arrow's weather predictions, the predictions estimated by the fitted model were as accurate as pulling predictions out of a hat. It is not surprising that first mover digital services collect large amounts of data. The academic community and managerial world praise it, the larger amount the better quality of predictions and statistical power. However, when it comes to using pre-existing data to understand customer satisfaction this does not seem to hold true. In the future, if relevant theories are developed to unite archive data with service quality and customer satisfaction measurements, the pre-existing data can become useful in this context.

5.1.2. The First Mover Disadvantage

Another explanation to the lack of explanation value could be the fact that Good Doctor Online was established in 2006, meaning nine- and twelve years prior to the studies, respectively, whereas the Swedish digital veterinary service was launched in 2016, only three years prior to this study. Moreover, the study by Yang et al., (2018), also measured customer satisfaction from first-time consultations as well as subsequent period consultations, indicating that Good Doctor Online is more established on the Chinese market, with customers who had used the service before and thereby had prior experience and expectations that their ratings could be based on. This is consistent with relevant customer satisfaction theories drawing on expectations. The data for the digital veterinary service on the other hand, showed that 89% of the customers were first time customers. This means that a great majority of customers have no prior experience of the service.

It can be argued that analyzing the consultations in multi-periods, as in Yang et al., (2018), could have given better results. However, that was not possible for the digital veterinary service, due to the high share of first-time customers. This indicates that there could be a first mover disadvantage when it comes to analyzing customer satisfaction. When there are no other actors offering a similar service on the market there are no comparisons for the customer to make. As the veterinarian mentioned during the interview, the only expectation customers have is that the service will be more effective than a regular consultation at the clinic, which it is since it does not require as much time, money and effort from the customers. When a second mover enters the market, it

may lead to the first mover's data becoming more valuable, since customers may be more sensitive, knowing that there is an alternative with the same efficiency.

5.1.3. Difficulties with Categorical Measurements of Customer Satisfaction

Perhaps the popular rating systems used today need alternatives. As seen in the data analysis, the most common way of measuring data - using OLS could not be used to analyze this data set. This due to the fact that the data was not normally distributed, most likely, since the dependent variable was categorical, using a scale of 1-5. This is very common in rating systems today, and customers are often asked to rate services on a Likert scale from 1-5 or 1-7. However, from a statistical perspective, analyzing categorical variables is not optimal due to difficulties in interpreting results from logistic regressions. For example, it is unlikely that the average marketing manager knows how to interpret the natural logarithm in the dependent variable. A possible alternative to examining antecedents of customers satisfaction could be to analyze the customers that chose to not use the service, analogically with the story about Wald and the missing bullet holes in *section 3.3*.

5.2. Potential Shortcomings

Inevitably, this thesis has its pros and cons. One weakness to take into account is that there are risks of using one company and extrapolate the results based on that to a broader phenomenon. Therefore, it is questionable whether this thesis can represent digital service providers in general. However, this is moderated when considering the exploratory approach of this thesis.

Further, another weakness is the limited access to the data. Therefore, the research method of this thesis had to be redesigned due to the nature of the data. The use of previous models can be questioned as the studied data set was not in the exact same format. The reason for the limited access is that this thesis uses archive data from a company and therefore does not have access to the database itself. Some data points were also limited due to protection of company sensitivity. However, this weakness should not be considered to impact the reliability of the findings of this thesis.

The dependent variable of the model in this thesis, *customer satisfaction*, only has two outcomes; *Satisfied* or *Not Satisfied*. The observations lacking a customer satisfaction rate were treated as *Not Satisfied*. This is a broad assumption and can therefore be questioned. However, the choice of this interpretation is derived from the assumptions made in the studies by Yang et al., (2015), and Yang et al., (2018). These studies have been published in *Decision Support Systems* and *International Journal of Production Economics*; two international research journals publishing quality peer reviewed research. One should, however, be aware that this could have had a significant impact on the results of the analysis.

In the scope of a Bachelor thesis, time- and content constraints is another weakness. The time available to conduct the study was constrained by a deadline and the content was constrained by a page count. However, this impact is acknowledged by encouraging future research.

There is always a risk of author bias due to cultural background or perspectives which can affect the thesis legitimacy. It is also possible that the authors are biased towards data and results that supports their arguments. These possible weaknesses were handled by choosing a quantitative method, appropriate handling of data using statistical literature, and through consistent and thorough consulting with the tutor of this thesis and other external reviewers.

5.3. Conclusion

To conclude the outcomes of this thesis, it has been identified that measuring customer satisfaction using pre-existing data for first mover businesses offering a digitalized service requires new theories and models adapted for the digital era, where customers do not have prior experience or expectations of a service.

To answer the theoretical question; "How can digital companies measure the antecedents of customer satisfaction using their data and available theories?" this thesis has found that digital companies in fact face difficulties in measuring the antecedents of customer satisfaction by using pre-existing data and conventional theories. The problem lies in the fact that these theories were developed to measure survey data. Further these conventional theories base customer satisfaction on expectations and prior experiences of a service, which unfortunately is not available for customers of a first mover offering a digital service and is not captured by archive data. Additionally, the study also found that the measures are not useful at all for making strategic decisions, as the data does not give any indications as to what makes a customer satisfied.

Managerial heuristics are still important but with the emergence of digital business environments the vast amounts of data available has a great value potential if it can be explained and exploited. However, this is proven difficult as there exists a disconnection between pre-existing data and conventional theories. There can also be a risk with putting too much trust into data, as first movers suffer the risk of getting results that are not valuable for the development of their business ideas. This is no critique against data-science models and statistical tools, but rather an encouragement for academics to develop models that can be used on pre-existing data making customer satisfaction analysis more efficient. Referring to the introduction, before these theories are developed the hopes are that this thesis can help companies avoid the pitfall experienced by Arrow's Commander General:

> The commanding general is well aware that the forecasts are no good. However, he needs them for planning purposes.

The superiors of Kenneth Arrow during World War II (Anderson & Griggs, 2018)

5.4. Future Study Proposals

Due to the time and content constraints of this thesis, a proposal for future study is to do the thesis study over a longer period of time. This would make it possible to assess the value of company data as a first mover digital service faces its first competitors. Another proposal is to conduct a longitudinal study. That study would observe the same customers over several periods of time in a longer total period. It would be possible to investigate the value development of company data when customers make more than one transaction to surveil how the relationship between the company and the customer evolves to affect customer satisfaction. Another interesting perspective would be to look at second movers to assess if they can better measure the antecedents of customer satisfaction due to the existence of adequate prior expectations of their service. Another topic to investigate is improving measures of e-customer satisfaction. To understand the area further, one could conduct studies assessing the populations who prefer traditional services over digital providers.

6. References

Agria Djurförsäkring. (2010). *När blir hunden vuxen?*. Retrieved at 2019 March 25 from https://www.agria.se/hund/artiklar/om-hund/nar-blir-hunden-vuxen/

Agria Djurförsäkring. (2017a). Allt fler hundar och katter i Sverige. Retrieved at 2019 March 25 from https://www.agria.se/pressrum/pressmeddelanden-2017/allt-fler-hundar-och-katter-i-sverige/

Agria Djurförsäkring. (2017b). Katten - populäraste sällskapsdjuret i Sverige. Retrieved at 2019 March 25 from https://www.agria.se/katt/artiklar/om-katt/katten---popularaste-sallskapsdjuret-i-sverige/

Alawneh, A., Al-Refai, H., & Batiha, K. (2013). Measuring user satisfaction from e-Government services: Lessons from Jordan. *Government Information Quarterly*, 30(3). 277-288. doi: 10.1016/j.giq.2013.03.001

Allison, P., (2013, February 13). What's the Best R-Squared for Logistic Regression? [Blog post] Retrieved from: https://statisticalhorizons.com/r2logistic

Almqvist, G. (2018). Uncertainty and Complexity in Predictions From Big Data: Why Managerial Heuristics Will Survive Datafication (Managing Digital Information, No 1). Stockholm: Stockholm School of Economics. Retrieved from https://www.hhs.se/contentassets/a3083bb76c384052b3f3f4c82236e38f/managing-digital-transformation-med-omslag.pdf

Anderson, E., & Sullivan, M. (1993). The Antecedents and Consequences of Customer Satisfaction for Firms. *Marketing Science, 12*(2), 125-143. Retrieved from http://www.jstor.org/stable/184036

Anderson, S., Pearo, L., & Widener, S. (2008). Drivers of Service Satisfaction: Linking Customer Satisfaction to the Service Concept and Customer Characteristics. *Journal of Service Research,* 10(4). doi: 10.1177/1094670508314575

Anderson, K. (2018, September 25). Planning Advice from Nobel Laureate Kenneth Arrow. *Advisor Perspectives.* Received from: https://www.advisorperspectives.com/

Armstrong, R. L. (1987). The Midpoint on a Five-Point Likert-Type Scale. Perceptual and Motor Skills, 64(2), 359–362.

Armstrong, J.S., Green, K.C., & Graefe A. (2015). Golden rule of forecasting: Be conservative. Journal of Business Research, 68(8), 1717-1731.

Bartlett, J. (2014,February 8), R squared in logistic regression [Blog post]. Retrieved from https://thestatsgeek.com/2014/02/08/r-squared-in-logistic-regression/?fbclid=IwAR0XpVdFQ3yd754jOovM4iuojq4alHH0x80nSNsananqC7KXUWh6EMH1e-E

Bjerling, J., & Ohlsson, J., (2010). En introduktion till logistisk regressionsanalys. Göteborg: Göteborgs Universitet

Bryman, A., & Bell, E (2015). *Business Research Methods* (4. ed.). Oxford: Oxford University Press

Byttner, K. (2019, May 3). "Struntprat att AI kan förgöra mänskligheten" menar professorn som jämför dagens algoritmer med en bakteries intelligens. *Veckans Affärer.* Retrieved from https://www.va.se/nyheter/2019/05/03/matteprofessorn-ingen-anledning-att-frukta-algoritmer/

Cardozo, R. (1965). An Experimental Study of Customer Effort, Expectation, and Satisfaction. *Journal of Marketing Research,2*(3), 244-249. doi:10.2307/3150182

Chu, P., Lee, G., & Chao, Y. (2012) Service Quality, Customer Satisfaction, Customer Trust, and Loyalty in an E-Banking Context. *Social Behaviour and Personality: an international journal,* 40(8). 1271-1283. doi: 10.2224/sbp.2012.40.8.1271

Daniel, E., Wilson, H., & Myers, A. (2002). Adoption of E-Commerce by SMEs in the UK Towards a Stage Model. *International Small Business Journal,* 40(3). 253–270.

Denrell, J., (2005). Selection Bias and the Perils of Benchmarking. *Harvard Business Review.*

D'Midio, T., Dorton, D., & Duncan, E. (2015, February). Service innovation in a digital world. *McKinsey Quarterly.*

Dreber, A., & Johanesson, M. (2018). Vilka forskningsresultat kan vi lita på?. *Ekonomisk debatt 2,* 17-28.

Duignan, B. (2018). Occam's razor. In *Encyclopaedia Britannica. Retrieved from:* https://www.britannica.com/topic/Occams-razor

Farris, P. Bendle, N. Pfeifer, P. Reibstein, D. (2015). Marketing Metrics: The Managers Guide to Measuring Marketing Performance.

Gandomi, A., & Haider, M. (2015) Beyond the hype: Big data concepts, methods, and analytics. International *Journal of Information Management,* 35(2). 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

Giesecke K., Liberali, G., Nazerzadeh, H., Shanthikumar, J.G., & Teo C.P., (2018). Call for papers - Management Science - Special Issue on Data-Driven Prescriptive Analytics. *Management Science, Volume 64,* 2972-2972.

GP Bullhound, (2016). *EUROPEAN UNICORNS 2016*. Retrieved from http://www.gpbullhound.com/wp-content/uploads/2016/06/GP-Bullhound-Research-European-Unicorns-2016-Survival-of-the-fittest.pdf

Helbing, D., Frey, B.S., Gigerenzer G., Hafen E., Hagner, M., Hofstetter, Y., Van den Hoven, J., Zicari, R.V., & Zwitter, A., (2017, February 25). Will Democracy Survive Big Data and Artificial Intelligence?. *Scientific American*. Retrieved from https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/

Helson, H., (1964). *Adaptation-level theory.* New York: Harper and Row.

IBM Big Data & Analytics Hub. Infographic: The Four V's of Big Data. Retrieved 2019 April 1 from https://www.ibmbigdatahub.com/infographic/four-vs-big-data

Investopedia, (2019). *First Mover Definition*. Retrieved from https://www.investopedia.com/terms/f/firstmover.asp

James, G., Witten, D., Hastie, T., & Tibshirani, R., (2017). *An Introduction to Statistical Learning*. New York: Springer Science+Business Media.

Kalakota, A.R., & Whinston A.B., (1997). *Electronic Commerce: A Manager's Guide*. Addison-Wesley Publishing Company, Inc.

Kamis, A., Yao, Y., Kim, S., (2014). An empirical validation of the patient-centered e-health framework in patient-focused websites. *CAIS 34*.

Khalifa, M., & Liu, V., (2003). Determinants of Satisfaction at Different Adoption Stages of Internet-Based Services. *Journal of the Association for Information Systems, Volume 4*, 206-232.

Laney, D., (2001). *Application Delivery Services*. Stamford: META Group Inc. Retrieved from https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Lee, G.G., & Lin, H.F., (2005). Customer perceptions of e-service quality in online shopping. *International Journal of Retail & Distribution Management, Volume 33*, 161-176.

Lennen Merckx, J., (2018). Allt fler träffar doktorn på nätet. *SVT Nyheter*. Retrieved from https://www.svt.se/nyheter/inrikes/allt-fler-traffar-doktorn-pa-natet

Lewicki, R.J., & Bies, R. (1998). TRUST AND DISTRUST: NEW RELATIONSHIPS AND REALITIES. *Academy of Management Review, Volume 23*, 438-458.

Lewicki, R.J., Tomlinson E.C., & Gillespie, N., (2006). Models of Interpersonal Trust Development: Theoretical Approaches, Empirical Evidence, and Future Directions. *Journal of Management, Volume 32*.

Lociacono, E., Watson, R.T., & Goodhue, D. (2000). *WebQual ™: A Web Site Quality Instrument,* (Working paper). Worcester: Polytechnic Institute.

Oliver, L.R. (1977). Effect of expectation and disconfirmation on post exposure product evaluations: an alternative interpretation. *Journal of Applied Psychology,* 480-486.

Oracle, (2013). Bringing the Value of Big Data to the Enterprise [Pamphlet]. Retrieved from http://www.oracle.com/us/products/database/big-data-appliance/value-of-big-data-brief-2008771.pdf

Parasuraman A, Zheitmal VA, Berry LL. (1985) SERVQUAL: a conceptual model of service quality and its implications for future research. J Mark. 49(1):41-50.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L., (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing, Volume 64*, 12-40.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L., (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research. *Journal of Marketing, Volume 58*, 111-124.

Parasuraman, A., & Grewal, D. (2000). The Impact of Technology on the Quality-Value-Loyalty Chain. *Journal of the Academy of Marketing Science, 28*, 168-174.

Persson, A., & Ryals, L. (2014). Making customer relationship decisions: Analytics v rules of thumb. *Journal of Business Research, Volume 67(Issue 8)*, 1573-1786.

Piete, J.D., Valverde, H., Marinec, N., Jantz, R., Kamis, K., Lazo de la Vega, C., Pinto, B., (2014). Establishing an independent mobile health program for chronic disease self-management support in Bolivia. *Front. Public Health 2.*

SAS Institute, (2019). What is Big Data?. Retrieved 2019 April 1 from https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Schwager, A., & Meyer C. (2007). Understanding Customer Experience. *Harvard Business Review*. Retrieved from https://hbr.org/2007/02/understanding-customer-experience

Wangenheim, F. (2003). Situational Characteristics as Moderators of the Satisfaction-Loyalty Link: An Investigation in a Business-to-Business Context. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour, Volume 16*.

Wolfinbarger, M.F., & Gilly M.C. (2002). *.comQ: Dimensionalizing, Measuring and Predicting Quality of the E-tail Experience* (Working paper). Cambridge: Marketing Science Institute. 02-100.

Wübben, M., & Wangenheim, F.V., (2008). Instant Customer Base Analysis: Managerial Heuristics Often "Get it Right". *Journal of Marketing, Volume 27*. 82-93.

Xu, J., & Quaddus M. (2009). *E-business in The 21st Century: Realities, Challenges And Outlook* (Volume 2). Singapore: World Scientific Publishing Co. Pte. Ltd.

Yang, H., Guo., X., & Wu, T., (2015). Exploring the influence of the online physician service delivery process on patient satisfaction. *Decision Support Systems, 78*, 113-121.

Yang, Y., Zhang, X., & Lee, P.K.C., (2018). Improving the effectiveness of online healthcare platforms: An empirical study of multi-period patient-doctor consultation data. *International Journal of Production Economics, Volume 207,* 70-80.

Yüksel, A., & Yüksel, F., (2008). *Tourist Satisfaction and Complaining Behavior: Measurement and Management, Issues in the Tourism and Hospitality Industry. Chapter 4: Consumer Satisfaction Theories: A critical review*. New York: Nova Science Publishers.

Zeithaml, V.A., Parasuraman, A., & Malhotra, A. (2005). E-S-QUAL A Multiple-Item Scale for Assessing Electronic Service Quality. *Journal of Service Research, Volume* 7. 2-21. doi:10.1177/1094670504271156

7. Appendix

Appendix 1. **Dimensions of WEBQUAL**

1) Informational Fit-to-Task
2) Interactivity
3) Trust
4) Response time
5) Ease of understanding
6) Intuitive operations
7) Visual Appeal
8) Innovativeness
9) Flow-Emotional Appeal
10) Consistent Image
11) On-Line Completeness
12) Better than Alternative Channels

(Lociacono et al., 2000)

Appendix 2. **Demographics for all consultations**

Variables	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Group 5**
Panel 1: Gender of dog owner					
Male	76%	75%	69%	78%	76%
Female	24%	25%	31%	22%	24%
Panel 2: Age of dog owner					
Children	0%				
Young adults	36%	36%	26%	29%	35%
Adults	30%	53%	38%	53%	52%
Middle-aged	28%	7%	30%	10%	8%
Seniors	5%	4%	7%	7%	6%
Panel 3: Age of dog					
Puppy	24%	25%	18%	23%	26%
Teenager	29%	30%	26%	29%	31%

Adult	19%	19%	20%	19%	19%
Middle-aged	18%	17%	21%	18%	16%
Senior	9%	9%	14%	10%	8%
Panel 4: Illness severity					
Other	11%	10%	10%	12%	
Eyes	8%	9%	10%	8%	
Toxic	3%	2%	2%	4%	
Urinal	3%	3%	7%	$%	
Respiratory	9%	10%	8%	9%	
Intestinal	22%	22%	15%	22%	
Genital	3%	3%	4%	4%	
Feeding & general care	0%	0%	0%	0%	
Lameness	6%	5%	5%	6%	
Skin, fur & ears	21%	23%	27%	18%	
Wounds	5%	5%	5%	5%	
Oral & dental	2%	2%	4%	2%	
Panel 5: Insurance					
Free consultations	96%	97%	96%	95%	99%
Paid consultation	4%	3%	4%	5%	1%
Panel 6: Type of customer					
First consultations	89%	89%	89%	90%	89%
Not first consultations	11%	11%	89%	10%	11%
Panel 7:Hours					
Regular hours	38%	38%	49%	37%	40%
Uncomfortable hours	62%	62%	51%	63%	60%
Panel 8: Waiting time					
Long waiting time	6%	5%	6%	6%	6%
Medium waiting time	26%	26%	26%	27%	26%
Short waiting time	68%	69%	68%	67%	68%
Panel 9: Duration					
Long duration	73%	76%	68%	69%	75%
Medium duration	24%	22%	26%	26%	22%

Short duration	3%	2%	6%	5%	3%

Group 1: All consultations, Group 2: Consultations with rating, Group 3: Consultation with rating below 4, Group 4: Consultation without rating, Model 5: Consultation for *Intestinal* and *Skin, Fur & Ears*