CATÓLICA LISBON
BUSINESS & ECONOMICS

Mobile Health Applications –

A Natural Language Processing Review Analysis

Maximilian Simmerer

Dissertation written under the supervision of
Prof. Henrique Martins

Dissertation submitted in partial fulfilment of the requirements for the MSc in Strategy & Entrepreneurship, at the Universidade Católica Portuguesa, 7th April 2021.

Abstract

Title: Mobile Health Applications – A Natural Language Processing Review Analysis

Author: Maximilian Simmerer

As smartphone use becomes ubiquitous, mobile health (mHealth) is becoming increasingly important. The COVID-19 pandemic has been linked to the increasing download rate of mHealth applications, as alternative healthcare solutions are sought by citizens restricted due to public health measures. Further evidence is necessary in order to prove if there was an increasing positive perception of healthcare applications by the population. This dissertation seeks to qualitatively and quantitatively evaluate this by studying the evolution of objective data like reviews and ratings of preselected mHealth applications in the Google Play Store during 2020. Further analysis focused on customer behavior is necessary to establish conditions and practices that support continuous growth and stability for mHealth providers. In order to analyze the mHealth applications development as well as the content of the reviews, natural language processing (NLP) tools were used. There is no published research or similar studies analyzing mHealth reviews with NLP tools for 2020. The results indicated no clear trend towards increasing positive perception of mHealth applications. Customer preferences were gathered and categorized by the 'Unified Theory of Acceptance and Usage of Technology' framework. The analyzed mHealth applications showed insights in form of different preferences in accordance with their respective subgroups. One main finding of this study was how the technical composition of mHealth applications can strongly influence the customer perspective. Furthermore, the study reveals missed opportunities if form of cooperation between regulators, application developers and insurances.

Keywords: Natural language processing, word and key feature analyses, mHealth, reviews, Google Play Store, ratings, COVID-19, applications, healthcare, unified theory of acceptance and use of technology

ítulo: Mobile Health Applications - Uma análise do processamento da linguagem natural

Autor: Maximilian Simmerer

À medida que o uso de smartphones se torna onipresente, a saúde móvel (mHealth) torna-se cada vez mais importante. A pandemia COVID-19 tem sido associada ao aumento da taxa de download de aplicações de mHealth, num contexto em que os cidadãos procuram soluções alternativas de acesso à saúde, mas falta evidência sobre se houve uma perceção cada vez mais positiva destas soluções. Esta dissertação pretende avaliar qualitativa e quantitativamente essa tendência, estudando a evolução de dados como avaliações e classificações de aplicativos de saúde móvel pré-selecionados na Google Play Store durante 2020. Uma análise mais aprofundada com foco no comportamento do cliente é necessária para estabelecer condições e práticas que apoiem o crescimento contínuo e a sustentabilidade dos fornecedores de mHealth. Para analisar o desenvolvimento das soluções bem como o conteúdo das críticas (reviews), foram utilizadas ferramentas de processamento de linguagem natural (PLN). Não há trabalhos publicados analisando críticas de aplicações mHealth estas ferramentas para 2020. Os resultados não indicaram uma tendência clara para o aumento de uma perceção positiva sobre aplicativos de saúde móvel. As preferências do cliente foram reunidas e categorizadas pela estrutura da "Teoria Unificada de Aceitação e Uso da Tecnologia". Diferentes subgrupos estiveram associados a diferentes perceções sobre os aplicativos mHealth analisados. Uma das principais conclusões deste estudo foi como a robustez técnica dos aplicativos de saúde móvel pode influenciar fortemente a perspetiva do cliente. A tese aponta oportunidades para a cooperação entre reguladores, criadores de soluções de saúde digital e seguradoras.

Palavras-chave: Natural language processing, mHealth,, análise de palavras, Google Play Store, classificações, COVID-19, aplicativos, saúde, teoria unificada de aceitação e uso de tecnologia

Contents

List of Figures

List of Tables

Acknowledgements

With this thesis an important chapter of my life is now coming to an end. During this time, I have been able to grow as a professional, but above all as a person. This would not have happened without the help of the people I would like to thank now. Foremost, I would like to thank my parents, my brother and my friends for all their support and understanding. I know that I can count on every one of you, for any upcoming challenges in my future life. Finally, I would like to express my deep and sincere gratitude to my supervisor Prof., Dr. Henrique Martins, who supervised and reviewed my master's thesis. His invaluable guidance and constructive criticism during the preparation of this thesis helped me to reach new depths in understanding digital healthcare. His vision, sincerity and motivation to the topic has inspired me to fully integrate healthcare into my career path.

List of Abbreviations

ACA – Affordable Care Act

BfArM – Bundesinstitut für Arzneimittel und Medizinprodukte

B2C – Business to Customer

CHIP – Children's Health Insurance Program

COVID-19 – Corona Virus Disease 2019

CSV – Comma Separated Values

GSM –Global System for Mobile Communications

GDP – Gross Domestic Product

GPRS – General Packet Radio Service

DiGA – Digitale Gesundheitsandwendungen

DVG – Digitale Versorgung Gesetz

FDA – Federal Institute for Drugs and Medical Devices

GER– Germany

GDP – Gross Domestic Product

GPRS – General Packet Radio Service

GSM –Global System for Mobile Communications

JASON – Java Script Object Notation

LDA– Latent Dirichlet Allocation

mHealth – mobile applications for smartphones and tablets, used for monitoring health status and/or improving the health outcomes of patients

NLP – Natrual Language Processing

PDA– Personal Digital Assistant

SHI – Statutory Health Insurance

UTAUT– Unified Theory of Acceptance and Usage of Technology

U.S.– United States

WiFi, – Wireless Fidelity

WiMAX – Worldwide Interoperability for Microwave Access

3G – Universal Mobile Telecommunications System

1 Introduction

Nearly half of healthcare emergency departments in the United States are underutilized under the current COVID-19 pandemic regarding non-COVID-19 patients (Wong, Hawkins, Langness., Murrell, Iris, Sammann, 2020), which is a phenomenon faced by healthcare systems of many countries around the world. For example, over 28 million operations globally were cancelled or postponed during the peak of the pandemic (COVIDSurg Collaborative, 2020), resulting in an immense backlog of elective operations while impacting the quality of life for many patients (The Lancet Rheumatology, 2021). There is growing evidence that one of the main reasons for this occurrence is that patients avoid emergency departments out of fear of contracting COVID-19, postponing emergency visits and elective operations, which in return leads to increased patient morbidity and mortality (Wong et al., 2020). Moreover, COVID-19-based restrictions in hospitals and private healthcare facilities reduced the predominantly face-to-face interactions between patients and healthcare professionals. Accordingly, the digital transformation of existing services was needed to meet the new challenges. To put it more simply, the digital transformation of the healthcare sector has changed from a 'can do' to a 'must do' and thus the current pandemic can be seen as an enabler of digitalization in the healthcare sector. This has resulted in a global readjustment towards new tools and solutions to alleviate patient anxiety through alternative solutions, providing platforms that enable patients to continue to benefit from the services of the healthcare sector, without violating government mobility restrictions. These readjustments were already enhanced by regulatory measurements taken by many governments to contain COVID-19, obliging the population to switch to digital media in both the private and public sectors, creating a unique environment for the growth of digital health solutions.

Recent legislation regarding digital health solutions by the Food and Drug Administration (FDA) or the Federal Institute for Drugs and Medical Devices (*Bundesinstitut für Arzneimittel und Medizinprodukte*, or BfArM), enabled large investments for digital health innovations and infrastructure, underline the trend towards digital health solutions (Kadaki, Patel & Shah, 2020). In other words, the COVID-19 crisis has led to extraordinarily rapid transformations of service delivery using telehealth technologies (Peek, Sujan, & Scott, 2020). One of the focal points here was the focus on mHealth applications as a digital health solution. The ever-increasing number of smartphone users, their enormous availability through all socio-economic classes, as well as the multiple opportunities they offer, have been reason enough for many countries to implement so-called COVID-19 warning applications at national level. The great

variety and opportunities of mobile applications were identified as one of the most fitting solutions to counter COVID-19 problems, resulting in a 30% increased download rate of mHeatlh applications in Germany and the U.S. in 2020 (Stewart, 2020). Doctors and patients have proved to be more open to trying digital healthcare products and services during the COVID-19 lockdown, when most direct face-to-face consultations were strongly limited (PWC, 2020). In particular, mHealth applications for medical and public health practice, supported by mobile devices such as mobile phones, patient monitoring devices and personal digital assistants (PDAs) (WHO, 2018), provide a fitting solution to the problem of increasing morbidity, mortality, and fear of infection with COVID-19 during medical visits, while resuming parts of the services of healthcare providers. With no end to the COVID-19 crisis in the foreseeable future, its impact on the mHealth market in terms of patient use, perceptions and acceptance cannot yet be conclusively assessed. With the momentum of the current situation, mHealth enthusiasts have a unique opportunity to better understand customer preferences, based on the huge amount of data. To keep a positive client attitude, however, in terms of perceptions and adoption, deeper insights into customer needs are essential.

The statistical analyses of rating scores of healthcare applications' evolution under pandemic times represent an opportunity for research. Moreover, the most important key features can be identified from the customer perspective. These findings have the potential to be used to improve services and technology acceptance and to optimize existing and future models/strategies for mHealth services' development. This in turn can increase the adoption rate of healthcare applications, which will have a positive impact on patients' health. The improvement of general perceptions regarding healthcare applications can foster and propel innovation in the healthcare sector. More patients will use healthcare applications, reducing face-to-face time with healthcare professionals. This may save costs and improve the quality of patients' health, while minimizing the risks associated with transmissible diseases such as COVID-19, other viruses and nosocomial infections. Finally, healthcare system players such as healthcare professionals, policymakers, regulators and insurance companies can benefit from solid information about the market development in mHealth and end-users' perceptions, which can be determined from detailed methodological analyses of available mHealth application ratings and reviews.

1.1 Problem Statement

Given the unique environment of 2020 under the COVID-19 pandemic, statistical evidence is necessary to determine whether there is an increasing adoption rate and positive perception of healthcare applications by the population. Following the results, an analysis focused on the needs of customers is necessary to establish the continuous growth and stability of mHealth applications. Existing research on this subject has not led to the publication of official results yet, taking the year 2020 and the usage of NLP into account.

Research Questions

This thesis aims to explore the evidence for a change in the perception of healthcare applications during the COVID-19 pandemic in 2020, and to provide corresponding recommendations for healthcare stakeholders based on customer preferences in the form of reviews, by a simple to adapt NLP taxonomy.

RQ1

Is there a change in perception by reviewers in the form of an increased or decreased rating score in the pandemic year 2020? Can cross-country and cross-sectoral differences be identified?

RQ2

What are the reasons for significantly increased or decreased rating scores for the analysed applications in 2020?

RQ3

What are the most important mHealth key features for customers based on the reviews? How should existing and future application developers position themselves, using NLP?

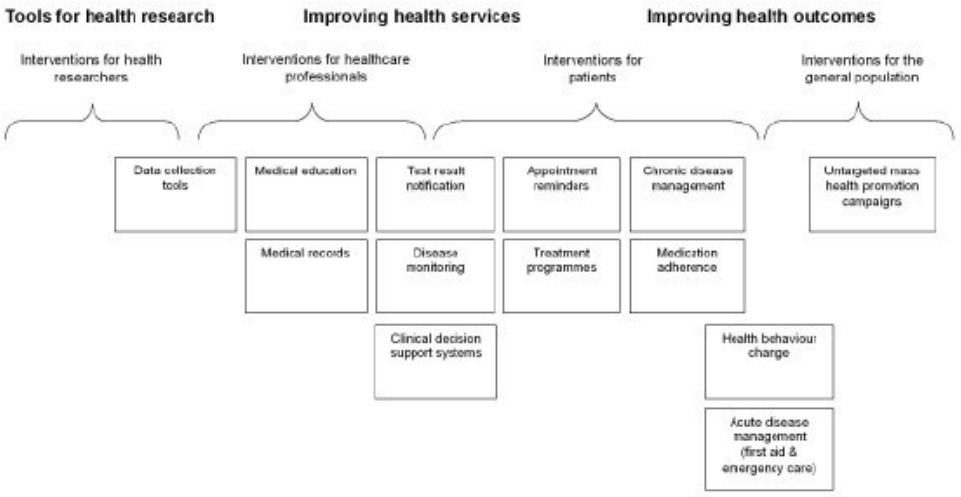
1.2 Mobile Health Applications

Before answering the research questions, it is necessary to define the meaning of mHealth, or, more specifically, medical applications, as many definitions and perspectives in the existing literature differ from each other, and to determine what classification should be used for this study. One of the first definitions of the concept (Istepanian et al., 2004, p.405) described mHealth as “mobile computing, medical sensor and communications technologies for

healthcare". In 2009, the United Nations and the Vodafone Foundation added wireless technologies to the definition (e.g. Bluetooth, GSM/GPRS/3G, WiFi, WiMAX), taking the transmission of various health-related data and services by mobile devices such as smartphones, mobile phones, PDAs, laptops and tablet personal computers into account. The World Health Organisation (WHO) defines mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices and personal digital assistants (PDAs)" (WHO, 2018, p.6). In 2010, Akter and Ray, in 2010 defined mHealth as a "personalized and interactive service whose main goal is to provide ubiquitous and universal access to medical advice and information to any users at any time" (Akter et al., 2010, p.75) over a mobile platform. Kumar et al. (2013) defined mobile health as application of mobile technology, which includes wireless devices and sensors, for monitoring health status or improving health outcomes. For the purpose of this dissertation, medical applications are defined as mobile applications for smartphones and tablets, used for monitoring health status and/or improving the health outcomes of patients.

The technological innovations of the past five years have resulted in a tremendous increase of applications for any kind of use, including facilitating the treatment of patients with chronic diseases. The online application stores are an example of such an innovation, being places where application providers can offer and sell their applications to end-users. The most important players in the market are the Google Play Store and the Apple App Store. A categorization of the applications is required, as mHealth applications can differ in their purpose and are not always considered medical applications. For this thesis, the following classification framework by Free, Phillips, Felix, Galli, Patel, Edwards (2010) was used, dividing mobile health applications into three main categories: Tools for health research, including interventions for health researchers; tools to improve health services, including interventions for healthcare professionals and interventions for patients; and tools to improve health outcomes, including interventions for patients and interventions for the general population (Free et al., 2010).

Figure 1: mHealth Application Classification



Source: From 'The effectiveness of M-health technologies for improving health and health services: a systematic review protocol,' by Free, Phillips, Felix, Galli, Patel, Edwards, 2010, *BMC research notes, 3,* p.4 (https://doi.org/10.1186/1756-0500-3-250). Copyright ResearchGate GmbH

Considering the already chosen definition of mHeatlh applications and the research questions of this study, applications representing interventions for patients were chosen as the application category/ type. Furthermore, it should be considered that mHealth is not used in isolation but rather is becoming, or already being, an interchangeable part of many healthcare systems, accelerating investments in recent years, especially under the current pandemic. For example, the German government approved the Act to Improve Healthcare Provision through Digitalization and Innovation (the Digital Healthcare Act or *Digitales-Versorgungs-Gesetz*, DVG), in September 2019 (Federal Ministry of Health, 2019), simplifying and reducing the legislative regulatory process of the BfArM, while making various investments in innovation, IT infrastructure and more, recently supported by the introduction of the law on the digital modernization of healthcare and nursing care in November 2020 (Kadakia, Patel & Shah, 2020). Also in the U.S., mHealth applications experienced recent regulatory changes, as many temporary policies to support digital health innovation during the pandemic were introduced by the FDA (Kadakia et al., 2020). In order to understand the role of mHealth applications in the current healthcare systems in the U.S. and Germany, it is necessary first to understand the broad healthcare systems of both countries.

1.3 Healthcare Systems in Germany and the U.S.

In order to understand the differences between the German and U.S. healthcare systems, it is important to have a look at the culture of each country first. The German constitution advocates the welfare state principle, which in some respects subordinates individual freedom in favour of the collective in order to protect individuals in need. This enables German citizens to claim benefit entitlements from the government in specific areas. U.S. citizens perceive their civil rights as natural, inalienable rights, not given or granted by any official institution. As a result, Germans think much more collectively, while Americans think far more individually. Translating this to the healthcare system/coverage of the two countries, Germans think of healthcare as a fundamental right that should be provided by the government. This is not the case for Americans. The individual freedom to choose one's personal destiny with only limited government influence is deeply embedded in American culture and Americans' perceptions on healthcare coverage. The adaptation of healthcare coverage to these fundamentally distinct perceptions results in different organizational healthcare structures and strategic approaches in the two countries. The structural differences are presented in the table below, considering the overall healthcare systems and their approaches to digital health.

Table 1: Country Comparison of Healthcare Systems

	Germany	**United States**
Universal Health Coverage	• The German system provides universal coverage for their citizens, along with a benefit package (Blümel and Busse, 2020). • "Health insurance is provided by two subsystems: statutory health insurance (SHI), consisting of competing, non-profit, nongovernmental health insurance plans known as sickness funds, and private health insurance"(Blümel et al., 2020, para. 3).	• The U.S. has no universal health insurance coverage. Almost 92% of the population has coverage in 2018, resulting in 27.5 million uninsured people (Tikkanen, Osborn Mossialos, Djordjevic and Wharton, 2020). • The health insurance is provided by multiple subsystems, of a private and public nature. • Employer-sponsored health insurance. • Individual-sponsored health insurance. • Medicate and Medicare. • Affordable Care Act (ACA). • Children's Health Insurance Program (CHIP). • TRICARE (Tikkanen et al., 2020)

The Role of the Government	• "The German health care system shares the decision-making powers among the federal and state governments and self-regulated organizations of payers and providers" (Blümel et al., 2020, para. 6). • "About 88% of the population receives primary coverage through sickness funds and 11% through private insurance. There were 109 sickness funds in January 2019" (Blümel et al., 2020, para. 13).	• The U.S. government has only a insignificant role in regards of directly owning and supplying providers (Tikkanen et al., 2020), with the following exceptions: • "Setting legislation and national strategies; • Administering and paying for the Medicare program; • Cofunding and setting basic requirements and regulations for the Medicaid program; • Funding health insurance for federal employees; • Regulating pharmaceutical products and medical devices; • Running federal marketplaces for private health insurance; and • Providing premium subsidies for private marketplace coverage"(Tikkanen et al., 2020, para. 13).
The Role of Public Health Insurance	• "In 2017, total health expenditures made up 11.5% of the gross domestic product (GDP). Of this health spending, 74% was publicly funded, and most of that spending (57%) went" (Blümel et al., 2020, para. 12) towards the SHI. • "Sickness funds are financed through compulsory wage contributions, levied as a percentage of gross wages up to a ceiling" (Blümel et al., 2020, para. 15).	• In 2017, public spending accounted for "approximately 8% of the GDP. Federal spending represented 28% of total healthcare spending. Federal taxes fund public insurance programs, such as Medicare, Medicaid, the CHIP, and military health insurance programs. The Centers for Medicare and Medicaid Services is the largest government source of health coverage funding" (Tikkanen et al., 2020, para. 17).
The Role of Private Health Insurance	• "In 2017, private health insurance accounted for 8.4% of total health expenditures" (Blümel et al., 2020, para. 18). • "This includes substitutive coverage purchased by individuals who are exempt from or can opt out of the SHI (such as higher-income individuals), as well as supplementary policies bought by	• "Spending on private health insurance accounted for one third (34%) of total health expenditures in 2018. Private insurance is the primary health coverage for two thirds of Americans (67%). The majority of private insurance (55%) is employer-sponsored, and a smaller share (11%) is purchased by individuals from for-profit and

	sickness fund enrollees" (Blümel et al., 2020, para. 18). • "In June 2018, there were 41 substitutive private health insurance companies in Germany, of which 25 were for profit" (Blümel et al., 2020, para. 19). • "The privately insured pay a risk-related premium, with separate premiums for dependents; risk is assessed only on entry, and contracts are based on lifetime underwriting" (Blümel et al., 2020, para. 20).	nonprofit carriers" (Tikkanen et al., 2020, para. 21). • "Most employers contract with private health plans to administer benefits. Most employer plans cover workers and their dependents, and the majority offer a choice of several plans. Both employers and employees typically contribute to premiums" (Tikkanen et al., 2020, para. 22).
Country Comparison of Digital Healthcare		
Strategic Approach	The Digital Healthcare Act (Digitale–Versorgung–Gesetz or DVG) introduced 2019 entitles those insured by the statutory health insurance provider to cover benefits for certain digital healthcare applications (bfarm, 2020). The applications have to fulfil specific requirements in order to be officially classified as digital health applications (Digitale Gesundheitsandwendungen, DiGA). The DiGA and BfArM are both controlling bodies supporting and regulating the process of digital applications. The applications cover the insured only after fulfilling the requirements.	Recent developments have led to the increased coverage of digitally provided services such as telehealth. The Medicare program introduced an initiative (Medicare Part B) which covers some telehealth services in 2019. During the pandemic, Medicare has expanded its telehealth coverage, temporarily allowing clinicians to provide services across state lines (chironhealth, 2020). This represents a historical change of regulations in the U.S., allowing more standardized processes across the country (Gerke et al.,2020; Kadakia et al., 2020). Still, this temporary change is only accepted for the duration of the pandemic. Post-COVID models still need to be developed, in order to maintain the status quo. Furthermore, a set of standards, including coverage rules, guidelines and other reimbursement strategies for digital solutions, is still missing.

Source: Adapted from bfarm, 2020; chironhealth 2020; Gerke et al., 2020; Tikkanen et al., 2020; Kadakia et al., 2020; Blümel et al., 2020

Conclusion

The comparison of the two healthcare systems indicates that the U.S. healthcare system has a highly fragmented structure in regards to healthcare coverage, focusing on the individual decisions of employers and employees. The accountability of insurance plans can be limited by the geographical location and hospitals, resulting in additional costs if the insured use services which are out of scope. Regulatory measurements in regards to digital solutions such as

telehealth are limited, allowing cross-state insurance usage of telehealth services for the duration of the pandemic. The German healthcare system is highly concentrated and standardized across its various stakeholders. This results in a high degree of bureaucracy, limiting quick adaptation and innovation. However, the 2019 introduced regulatory measurements for digital healthcare and digital health applications enforce quality standards, innovation and investments for mHealth applications (Gerke et al.,2020; Kadakia et al., 2020). With this, a foundation for future digital health solutions was build. Considering the early stage of this development, statistical evidence cannot be obtained yet, as applications falling under the DiGA and the BfArM have not yet matured and usage is still on a small scale.

2 Methodology

This research study conducts an empirical data analysis of Google Play reviews of healthcare applications based on natural language processing (NLP) machine learning tools, providing statistical evidence on the increase/decrease in usage, or maintained usage, and positive or negative perceptions of mHealth applications, based on the customer reviews. The provided framework of NLP usage presents a simple way for researchers and application developers to identify key problems of offered services.

2.1 Concepts and Methods

For this study, online reviews from the Google Play Store were used. In general, the studied online reviews can be defined as opinions and/or experience of used services, written by consumers (europa.eu, 2015). The Google Play Store was chosen as platform to collect the reviews, based on the simplicity of collecting data using a laptop, and because it is one of the largest platforms for mHealth application reviews in the world. The following typology is used for mHealth reviews: A star rating from 1 to 5 is assigned, and a short text is uploaded by a reviewer, who provides a written opinion and/or relates experiences. Studying these reviews has the advantage of having access to vast amounts of data, not available through manual collection.

As the content of reviews is organized alphabetically, NLP tools were used in order to transform the content into numerical data. In addition, recommendations are provided for action based on the review insights, focusing on identifying unusual occurrences in the reviews and ratings analyses. Furthermore, the actual content of the studied reviews is explore more in-depth and the frequency of used words is used to gain a better understanding of patient needs and wants,

using simple NLP tools which can be used and adapted without any extensive background in coding.

2.1.1 Natural Language Processing

Natural language processing, also called computational linguistics, is the combination of AI and linguistics which allows processing, analysing and understanding large amounts of text data (Lee, 2019). In other words, NLP enables computers to read text, interpret it and measure and determine which parts are most relevant. It resolves "language ambiguity and adds useful numeric structure to data for many downstream applications"(Ilias, Kanavos and Adamopoulous, 2020, p.23774) . Considering the high value of data provided with each review, this study uses NLP tools to classify and structure the raw data, analyse the texts and interpret the outcomes of a vast number of reviews.

2.1.2 Unified Theory of Acceptance and Usage of Technology Model

After large amounts of data have been processed and structured using NLP tools, it is necessary to summarize these according to logical key characteristics. One of the most fitting models for the main research questions of this thesis is the unified theory of acceptance and usage of technology (UTAUT) model. The UTAUT model focuses on the technology adoption of consumers. It integrates several theories of technology adoption and provides a comprehensive view of factors influencing consumers' adoption behaviour (Ndayizigamiye Kante & Shingwenyana, 2020). As this thesis aims to identify customers' perceptions regarding the most important/influencing factors of mHealth applications, the UTAUT framework and its factors needed to be slightly adjusted in order to be ideally suited for this purpose. Instead of using these factors to explain why consumers adopt mHealth services in the first place, they are used as a basis to understand whether the services adopted match consumers' expectations, comprehensively explaining users' preferences regarding adopted mHelath services. Accordingly, the analysed content was divided into four categories using the UTAUT model's predefined factors.

2.2 Methodological Approach

The structure for empirical analyses was divided into seven steps, where each step represents the basis for the following step. The analyses were mainly conducted using Python, following

personally set rules and requirements in accordance with the general norms of data analysis. The steps for answering the research questions were as follows:

1. Decision-Making process for applications;
2. Data collection;
3. Review rating score and number of analyses;
4. Identification of outliers in 2020;
5. Text analyses of outliers in 2020;
6. Text analyses of all collected applications;
7. Key features' analyses; and
8. Recommendations.

2.3 Decision-Making Process for Applications in this Study

Taking the vast market of mobile applications into account, it was important to define the segment, scope, maturity and place of this study, which have a direct impact on the decision-making process. For the analysis of the applications, the study focuses on the countries Germany and the U.S., to make a direct comparison of the impact of COVID-19 between North America and Europe. In addition, a cross-industry comparison of financial applications was conducted, as it helps to understand whether the development is unique for this market segment or whether it applies to numerous market segments. As the platform for research, the Google Play Store was chosen. In order to answer the research questions, this study analyses ten healthcare applications from Germany, ten healthcare applications from the U.S. and ten financial services applications, for a cross-sectoral analysis. Furthermore, each application had to have a minimum of 3,000 reviews and a maturity of at least three years. The defined scope and maturity were based on the statistical significance each application has to deliver in order to represent statistical evidence while answering the stated research questions. To be more specific, the investigated applications were supposed to have at least 15 reviews per month over the past three years to circumvent bias in the average rating-per-month analysis. Because of these restrictions, the number of potential mHealth applications were extremely limited, especially for the German mobile application market.

2.4 Data Collection

For the data collection, it was necessary to convert the application and information into data frames and save these to comma-separated values (CSV) files. Therefore the necessary packages were installed. The applications were then assigned to app packages, followed by information scrapping for each individual application. With the info function, one can adapt the data set to the language and country one wants to analyse. To be more precise, a helper function for JSON objects was added, in order to get all valuable information of each application. As a next step, all available data in the determined language and country were scrapped from the chosen application. In order to ensure all reviews were collected, the cluster 'rating' and a subset 'newest' were included in the review scrapping. Finally, the subset category with 'newest' was dropped for a clean data set without any duplicates. For each application, a data frame was constructed and saved as a .csv file, including the following columns: UserName, Comment, Score, ThumpsUpCount, AppVersion, Timestamp, Reply content, Sort order and Review ID.

2.5 Review Rating Score and Number of Analyses

For this part, the newly created data frames were analysed in terms of the average rating score per quarter of a year and the number of reviews per quarter of a year. An additional analysis of ratings per month and number of reviews per month was performed to identify unusual occurrences and in order to fulfil the requirement of having at least15 ratings per month. The results were summarized in terms of their respective application segments and types, starting with the U.S. and followed by German mHealth applications and financial services applications. The identified outliers were further examined in the next step, particularly investigating each specific timespan when the unusual occurrences happened.

2.6 Identification of Outliers

After reviewing the rating scores and number of analyses, each subgroup was analysed for particularly conspicuous occurrences during the year 2020. For each significant outlier, a comparable counterpart was chosen in order to draw a comparison. The outliers were defined by either a very strong increase in the review growth rate and average rating development, or a very strong decrease in the review growth rate and average rating development, followed by a distribution analysis of the rating scores per month. Reviews with a rating of 5 were marked as 'VeryPositive', with 4 as 'Positive', with 3 as 'Neutral', with 2 as 'Negative' and with 1 as

‘VeryNegative’. The phases with a very high number of ‘Positive’ and ‘VeryPostive’ reviews and the phases with a very high number of ‘Negative’ and ‘VeryNegative’ reviews were split from the data frame, to do a phase-specific analysis, in order to identify the reasons.

2.7 Text Analyses of Outliers

For the identification of key reasons for either a major drop or increase in average rating scores, the exact timespans in 2020 were identified, extracted and analysed for the selected applications. For the word analyses of the reviews, some data wrangling was necessary. First, all reviews were grammatically corrected by an automated spelling corrector, to ensure meaningful and valuable reviews. Second, all non-alphabetic characters were removed and all words were converted to lowercase, because we wanted to use the words as features and it would have resulted in a frequency bias as, for example, the word ‘Go’ at the beginning of a review would count as a different word to ‘go’ in the middle of a sentence. This was followed by sentence tokenization, which divided comment text paragraphs into tokenized sentences. Next all stop words, such as ‘the’, ‘she’, ‘he’, ‘a’ et cetera were removed, as they had no value for the analyses.

Finally, the lemmatization method for normalizing the words of a sentence into their base root forms was used. For example, words such as ‘love’, ‘loving’ and ‘lovable’ have their origin in a single root word, which is ‘love’. The lemmatization method extracts the endings and beginnings of words, based on a list of common prefixes and suffixes (Kalla & Samiuddin, 2020). Contrary to the stemming method, the lemmatization method results in a base word instead of a stem, while with the stemming method the endings of words are dropped or slightly changed. For example, the terms ‘studied’, ‘studies’, and studying’ are all variations of the base word ‘manage’. A stemmer converts all of these terms to ‘stud’, so that they share the same base (Jabeen, 2018). For this analysis, the stemming method was not chosen, as it does not generate real words. For the German stop words, a small adjustment were done, in order to erase unnecessary words such as ‘der’, ‘die’, ‘das’ et cetera.

As a next step, the reviews were divided into positive and negative reviews. Using n-grams, the already tokenized sentences of each review were broken down into chunks of words. With a frequency counter, the top 20 occurring words from the positive and the negative reviews were generated in the form of a unigram (one word). In order to achieve the maximum output, the positive and negative reviews were also analysed in terms of binary and tertiary word

combination frequencies. The results were entered into a table, providing insight into the most frequently used words in unigrams, bigrams and trigrams for the respective subset data frame of each chosen application. In addition, a manual analysis was conducted of words and word combinations with regard to the greatest added value to answer the second research question. Words such as 'NameOfTheApp', 'app' and 'health' have a natural high frequency in the reviews, but they are not relevant for the identification of the reason why certain apps performed very well or very bad during a specific timespan. Therefore these words were not considered for the following analyses.

2.8 Text Analyses of All Collected Applications

Considering the insights gained from the text analyses of the identified outliers, a holistic text analysis of the defined subgroups of each segment was performed, in order to understand the reasons responsible for either positive or negative reviews, in their respective types. Therefore the data frames of each application in their respective subgroup were merged into one. For the holistic text analysis itself, the same procedure as with the outliers was performed, with the difference that the top 40 positive and negative results were examined in order to generate a overview without bias.

2.9 Key Features of Analyses and Recommendations

As explained in section 2.1.2, the UTAUT model was used for the text analyses of the application segments, identifying the most dominant customer preferences for each subgroup. The key features/factors of the UTAUT framework in relation to online reviews need to be understood as the features/factors responsible for the continued use of the applications.

Performance Expectancy

Performance expectancy is related to users' perceptions of how mHealth applications will improve their personal healthcare (Alaiad, Alsharo & Alnsour et al., 2019). This implies the quality of the outcome and the usefulness of the application features.

Effort Expectancy

Effort expectancy is related to the perceived simplicity and convenience of mHealth applications (A. Alaiad et al., 2019), focusing on the ease of use, the simplicity of the interface and the complexity of the offered services.

Social Influence

Social influence is related to the effect of customers' surroundings when they are making health decisions (A. Alaiad et al., 2019). The existing literature argues that patients seek guidance from their surroundings when making health decisions (Hoque et al, 2017). To be more precise, reviews containing words such as 'recommendation', 'recommended', 'advised' et cetera fall under the umbrella of social influence.

Facilitating Condition

Facilitating conditions are related to the existence of organizational and technical infrastructure which supports mHealth applications and services (A. Alaiad et al., 2019). Given the size of the Google Play Store platform and the wide variety of types of services and applications, the feature 'facilitating conditions' is not viable for a word frequency analysis. Instead of facilitating conditions, a new, more convenient key feature was developed. Given the technical complexity some applications embody, focusing on seamless transitions and processes was more valuable.

Process Expectancy

Process expectancy is related to the customer's perceived satisfaction level, regrading technical processes and functionality. In the key feature analysis, a personalized dictionary was designed for each subgroup, encompassing the properties of the key features mentioned above. Using the cleaned and merged data sets from the text analyses, the distribution of the key features was visualized in the form of a stacked barplot, over the complete lifespan of each subgroup. In addition, spider diagrams were created for a clearer understanding of customer preferences in each subgroup. One spider diagram is used to illustrate the general perceptions of the applications in the form of ratings. A second one is used to highlight customer preferences in a comprehensive way. The results will be used for a final recommendation.

3 Results and Analysis

This chapter covers the results of the stated analyses in chronological order. At the same time, an analysis of the results for each of the three main research questions will be presented throughout this chapter.

3.1 mHealth Application Data Sets

For the key feature analyses, each application will be clustered in the following order: By country, type and review development. Each Country has at least two different application types, which were generally summarized as in the following tables.

Table 2: U.S. mHealth Applications

App Name:	Type:	Number of Reviews:	Average Rating:	Maturity:
Ada	Medicine, Monitor	60104	4.7	4 years
WebMD: Check Symptoms	Medicine, Monitor	12354	4.4	10 years
mySugr	Medicine, Monitor	12704	4.6	7 years
Doc On Demand	TeleHealth	20520	4.9	7 years
K-Health Telehealth, Primary Care & Pediatrics	TeleHealth	4538	4.0	3 years
MyChart	TeleHealth	14388	4.2	3 years
Folllow My Health	TeleHealth	50263	4.5	9 years
Samsung Health	Health & Fitness	79078	3.9	6 years
Huawei Health	Health & Fitness	41342	4.3	4 years
Google Fit	Health & Fitness	57879	3.9	6 years

Source: Own creation

Table 3: German mHealth Applications

App Name:	Type:	Number of Reviews:	Average Rating:	Maturity:
Ada	Medicine, Monitor	6804	4.7	4 years

Tabletten Erinnerung für Medikamente und Pillen	Medicine, Monitor	4443	4.8	6 years
Menstruations-Kalender Clue: Perioden & Zyklus-App	Medicine, Monitor	8047	4.7	6 years
Menstruations-Kalender	Medicine, Monitor	36717	4.9	9 years
Fabulous – Motivierend!	Health & Fitness	3414	4.5	7 years
Huawei Health	Health & Fitness	8960	4.3	4 years
Fitbit	Health & Fitness	26968	3.7	9 years
Samsung Health	Health & Fitness	23393	3.9	6 years
Beurer ealthManager	Health & Fitness	5039	2.8	7 years
Google Fit	Health & Fitness	7818	3.9	6 years

Source: Own creation

Table 4: Finance Applications

App Name:	**Type:**	**Number of Reviews:**	**Average Rating:**	**Maturity:**
GooglePay: A safe & helpful way to manage money	Finance, Transaction Service	143230	4.0	3 years
CashApp	Finance, Transaction Service	118408	4.2	7 years
PayPal	Finance, Transaction Service	374530	4.2	12 years
WorldRemit	Finance, Transaction Service	31601	4.2	7 years
Acorns - Invest Spare Change	Finance, Trading	55897	4.4	6 years
Fidelity Investments	Finance, Trading	33271	4.3	11 years

Stash: Invest & Build Wealth	Finance, Trading	25064	3.7	5 years
TrueBill Budget Planner, Bill Tracker and Reminder	Finance, Budget Planner	5341	4.3	4 years
Mint: Budget, Bills, Finance & Tax Refund Tracker	Finance, Budget Planner	59304	4.5	11 years
Goodbudget: Budget & Finance	Finance, Budget Planner	6715	4.4	12 years

Source: Own creation

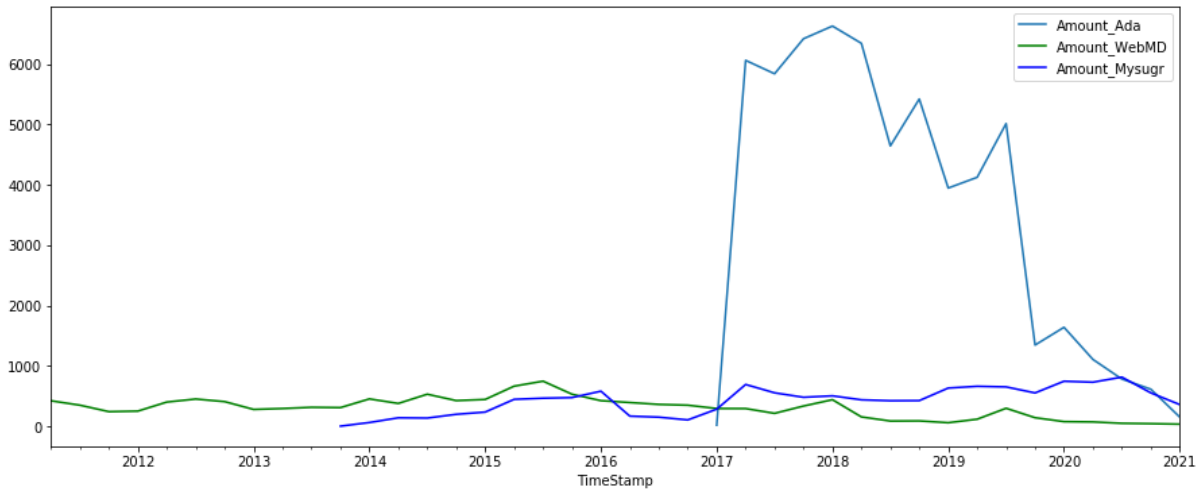
The focus of this study is on mHealth applications. These were grouped into the following types: 'Medicine', 'Monitor', 'Telehealth' and 'Health & Fitness', representing application for patient intervention. The rules regarding scope and maturity did not allow any German mHealth applications for the type 'Telehealth Monitor' to be further analysed. However, this type has a strong COVID-19 correlation and was therefore used for the U.S. mHealth applications.

3.2 Review of Data Collection

Summing up the cleaned results of the data collection, the group mHealth applications in the U.S. provided 353,170 reviews for further analysis. The group mHealth applications in Germany provided 124,799 reviews and the group financial applications in the United States provided 853,361 reviews. Accordingly, more than 1.3 million reviews were analysed for this study, of which approximately one third were related to healthcare. In addition, each group was divided into subgroups according to their respective types, creating similar data sets with regard to composition and structure.

3.3 Results: Research Question 1

Figure 2: U.S. mHealth Applications: Medicine, Monitor, Number of Reviews



Source: Own creation

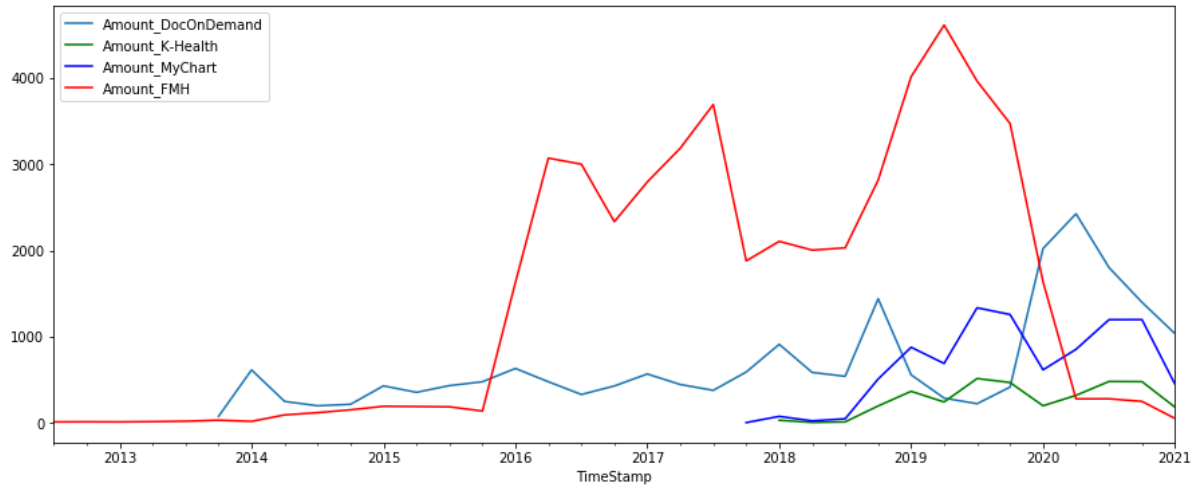
The segment mHealth applications in the U.S. for the types 'Medicine' and 'Monitor' indicates an increasing number of reviews for 2020 for the application 'mySugr', while 'Ada' experienced the most dominant phase in 2017. The third quarter of 2019 was marked by a significant decrease in reviews. Although the last quarter of 2019 indicates a recovering phase in the form of more reviews, a downward trend for 2020 can be identified. The application 'WebMD' and 'mySugr' indicate a stable number of reviews for 2020. Neither a significant increase nor decrease can be identified.

Figure 3: U.S. mHealth Applications: Medicine, Monitor, Rating Development

Source: Own creation

The figure 3 indicates a positive rating development for ‘Ada’ in 2020, while ‘Mysugr’ represents a slight downward trend, excluding the first quarter of 2020. One has to consider the high rating average of this ’Ada’ and ‘MySugr’, as during their entire lifespan the average monthly ratings for these applications did not fell below 3.0, which indicates overall high satisfaction. The application ‘WebMD’ shows a significant drop in the average ratings score over the whole of 2020. Nevertheless, based on the small number of reviews available for 2020, ‘WebMD’ most likely represents a stable positive/negative ratings score. A more in-depth analysis should be conducted to ensure there were no exceptional occurrences.

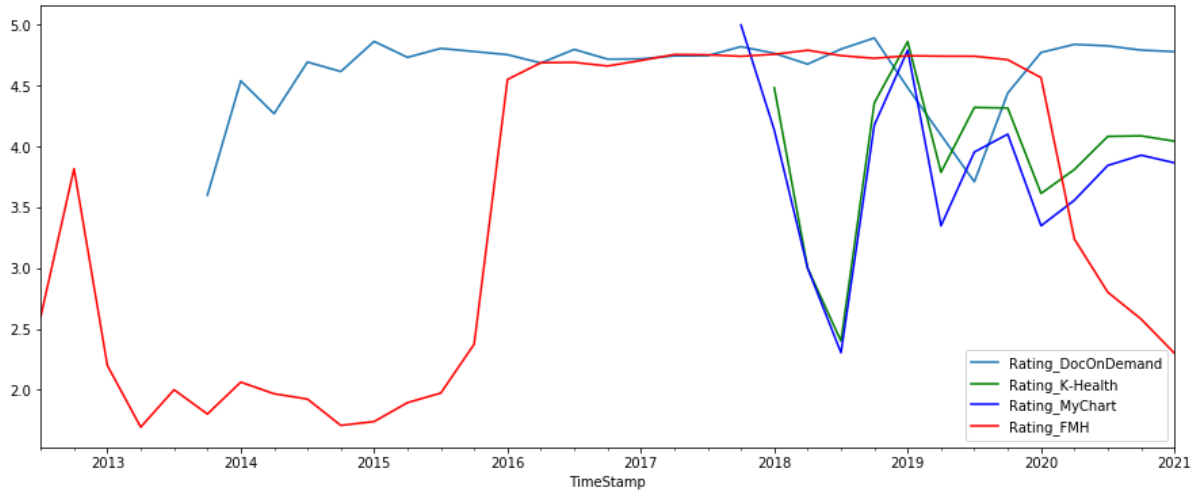
Figure 4: U.S. mHealth Applications: Telehealth Number of Reviews



Source: Own creation

The segment mHealth applications in the U.S. for the type ‘Telehealth’ indicate an increasing number of reviews in 2020 for the applications ‘DocOnDemand’, ‘K-Health’ and ‘MyChart’, representing the most dominant phase in the lifespans of these applications. The application ‘FollowMyHealth’ indicates an exceptional drop in the number of reviews for the year 2020.

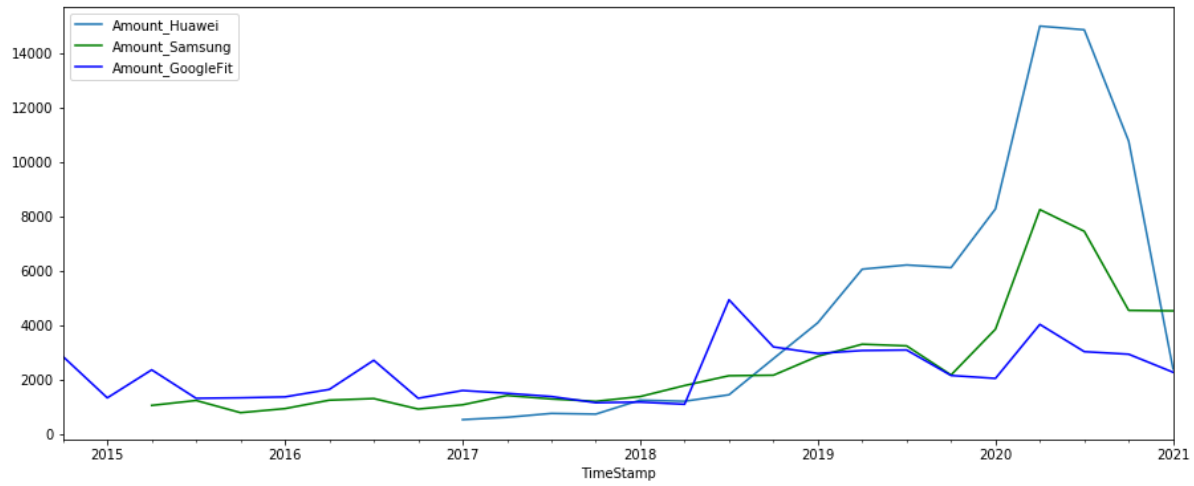
Figure 5: U.S. mHealth Applications: Telehealth Rating Development



Source: Own creation

The figure 5 indicates a positive rating development for 'DocOnDemand', 'K-Health' and 'MyChart' in the first three quarters of 2020 and a slightly stagnating development for the last quarter of the year. 'FollowMyHealth' represents an unusual exception, with a significant drop in number of reviews and rating scores. As the drop in number of reviews goes hand in hand with 'FMH's' decreasing ratings, an additional in-depth analysis should be conducted.

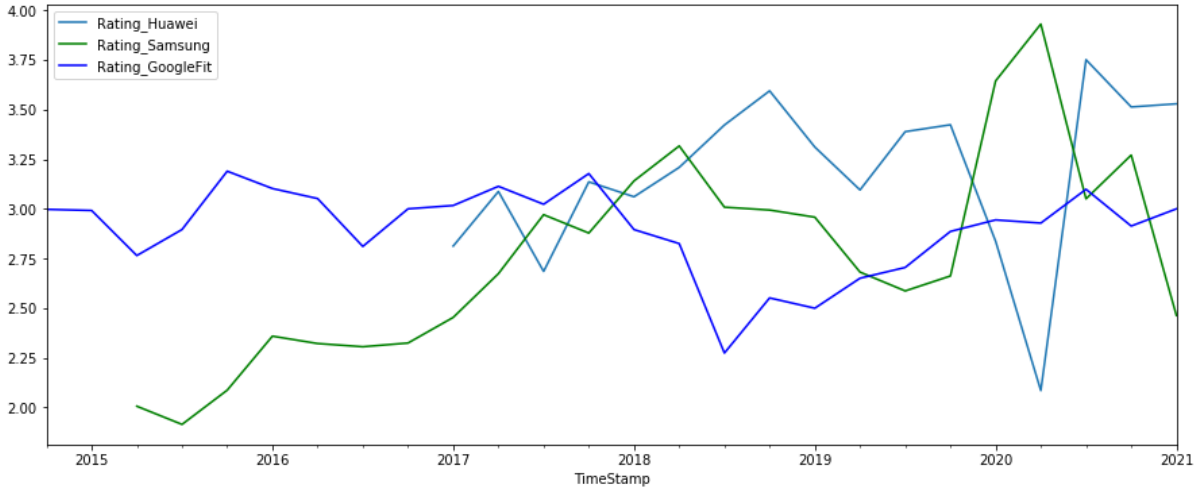
Figure 6: U.S. mHealth Applications: Health & Fitness, Number of Reviews



Source: Own creation

The segment mHealth applications in the U.S. for the type 'Health & Fitness' indicates an increasing number of reviews in 2020 for the applications 'HuaweiHealth' and 'SamsungHealth', both peaking in the first second and second quarters of 2020. The application 'Google Fit' indicates stable growth in the number of reviews for 2020.

Figure 7: U.S. mHealth Applications: Health & Fitness, Rating Development

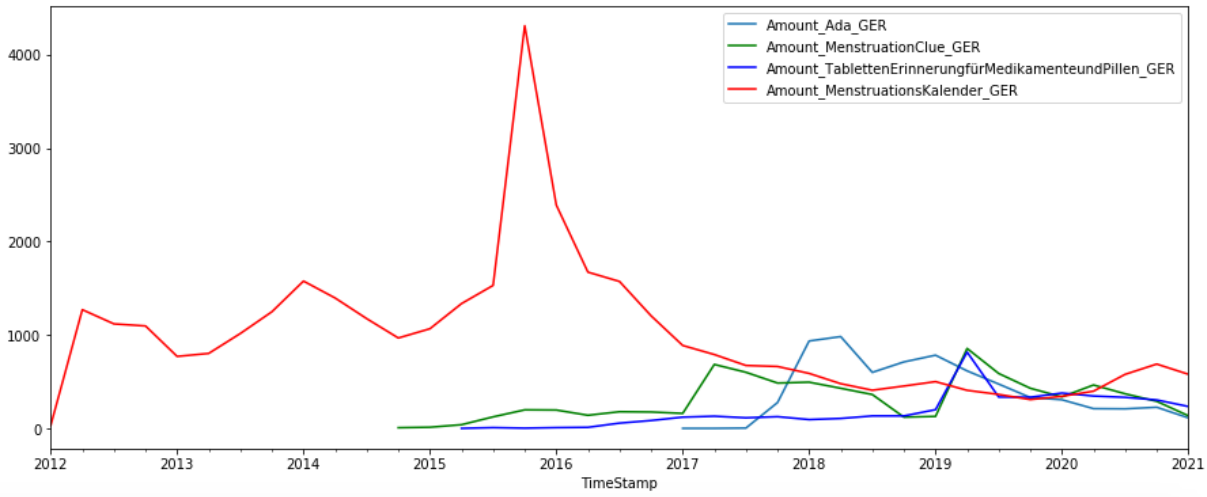


Source: Own creation

The figure 7 indicates a positive rating development for 'SamsungHealth' in the first quarter of 2020, followed by a significant drop for the rest of the year. By comparison, 'HuaweiHealth' started the year with an unusually sharp drop in average ratings, followed by a strong recovery for the rest of the year, reaching its all-time high in the third quarter of the year. The application 'Google Fit' experienced increasing rating development throughout the whole of 2020.

The next section presents the rating development and number of reviews of the German mHealth applications.

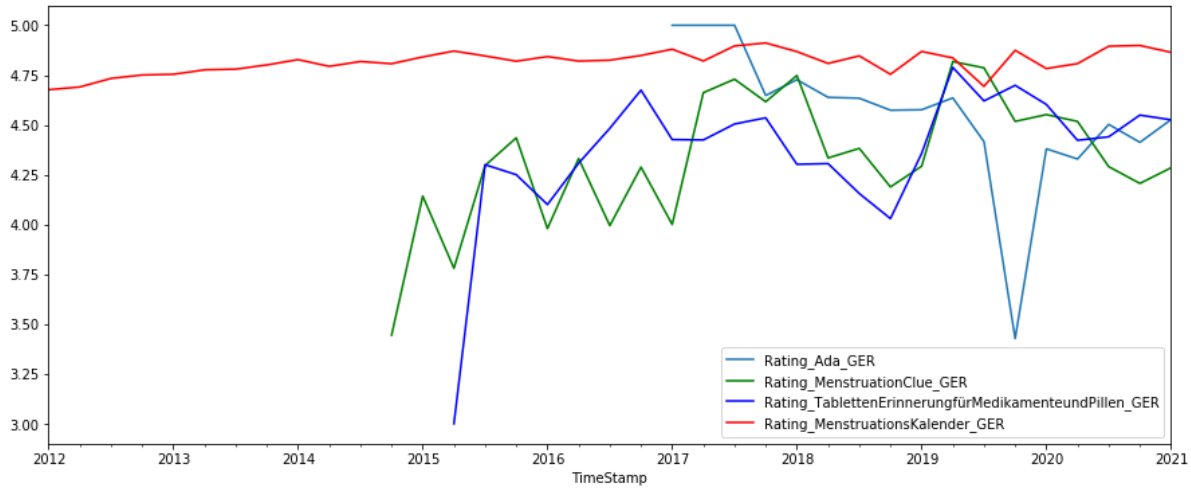
Figure 8: German mHealth Applications: Medicine, Monitor, Number of Reviews



Source: Own creation

In the segment mHealth applications in Germany for the types ‘Medicine’ and ‘Monitor’ all applications indicate slightly decreasing numbers of reviews in 2020, except for ‘MenstruationsKalender’.

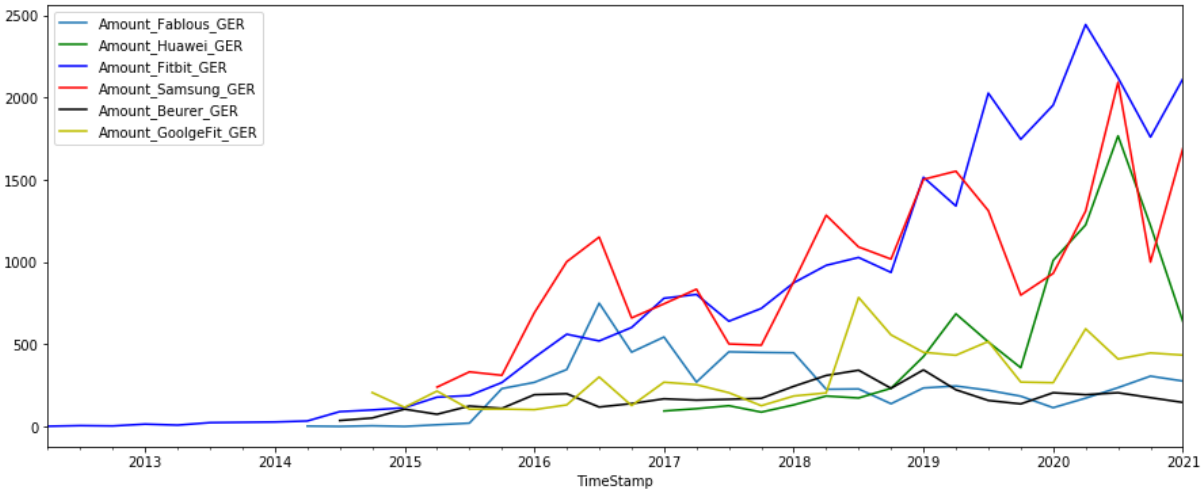
Figure 9: German mHealth Applications: Medicine, Monitor, Rating Development



Source: Own creation

Contrary to the development of the number of reviews, three out of the four applications indicate positive rating development, with the exception of ‘Menstruations-Kalender Clue: Perioden & Zyklus-App’, for which only a decrease in ratings averages could be identified. Still, the rating of 2020 is above the average rating for the lifespan of ‘Menstruations-Kalender Clue: Perioden & Zyklus-App’ and none of the ratings of any applications fall below 4.0 in 2020, indicating predominantly positive reviews.

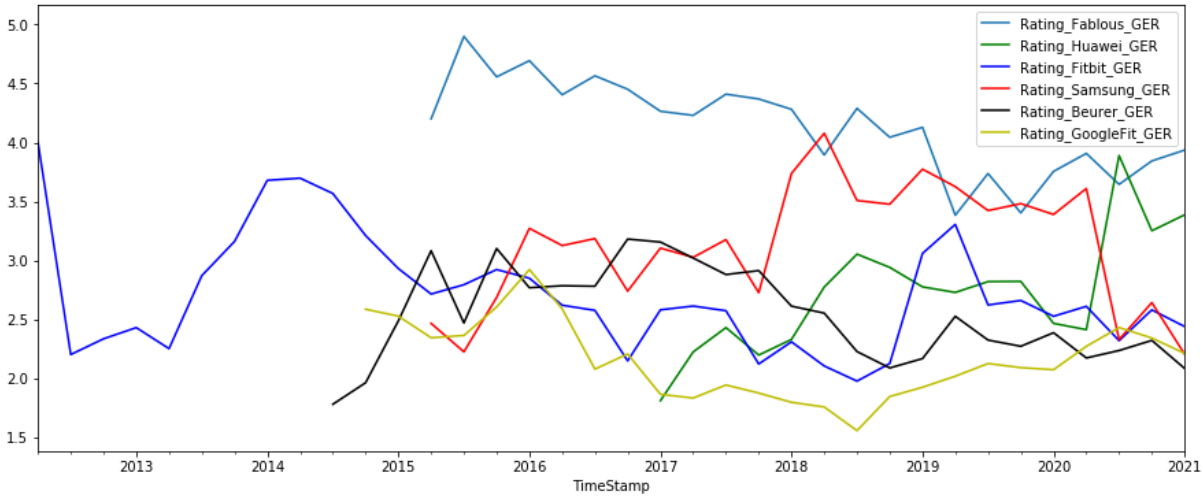
Figure 10: Germany mHealth Applications: Health & Fitness, Number of Reviews



Source: Own creation

In the segment mHealth applications in Germany for the type ‘Health & Fitness’, the applications ‘Huawei Health’ and ‘Samsung Health’ indicate, like their counterparts in the U.S., a strong increase in the number of reviews in the first two quarters of 2020. The application ‘Fitbit’ is highlighted by a strong increase in reviews over the past few years, indicating a drop in the number of reviews for the first two quarters over the past three years, followed by a strong recovery. The remaining applications indicate stable numbers of reviews with no significant drops or increases.

Figure 11: Germany mHealth Applications: Health & Fitness, Rating Development

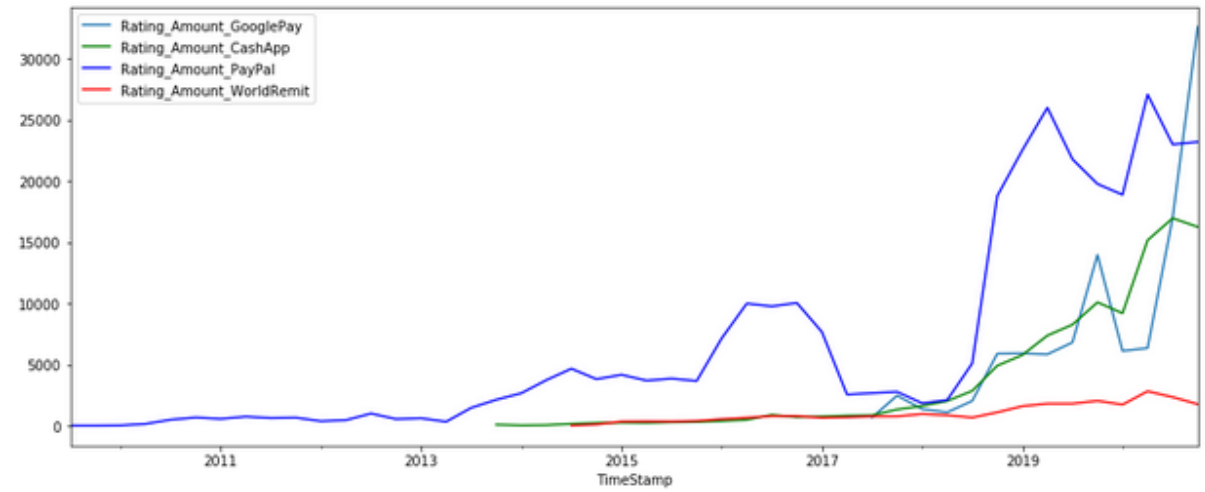


Source: Own creation

The figure 11 indicates a positive or stable rating development for the applications ‘Fitbit’, ‘Fablous’, ‘Google Fit’ and ‘Beurer’. By comparison, their counterpart ‘HuaweiHealth’ indicates a positive rating development in 2020, while ‘SamsungHealth’ indicates a negative rating development.

The next section presents the rating development and number of reviews of financial applications.

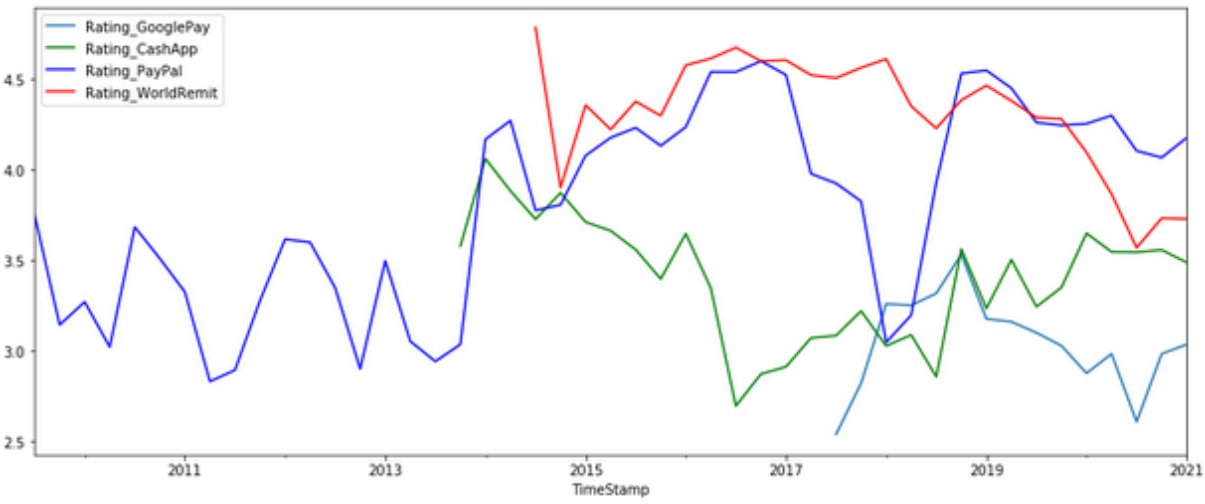
Figure 12: Finance Applications: Transaction, Number of Reviews



Source: Own creation

In the segment financial applications for the type 'Transaction', two thirds of the listed applications indicate a significant increase in the number of reviews in the first two quarters of 2020. While 'GooglePay' experienced strong growth throughout the whole year, the applications 'PayPal', 'World Remit' and 'CashApp' indicate stable development in the last quarter of the year.

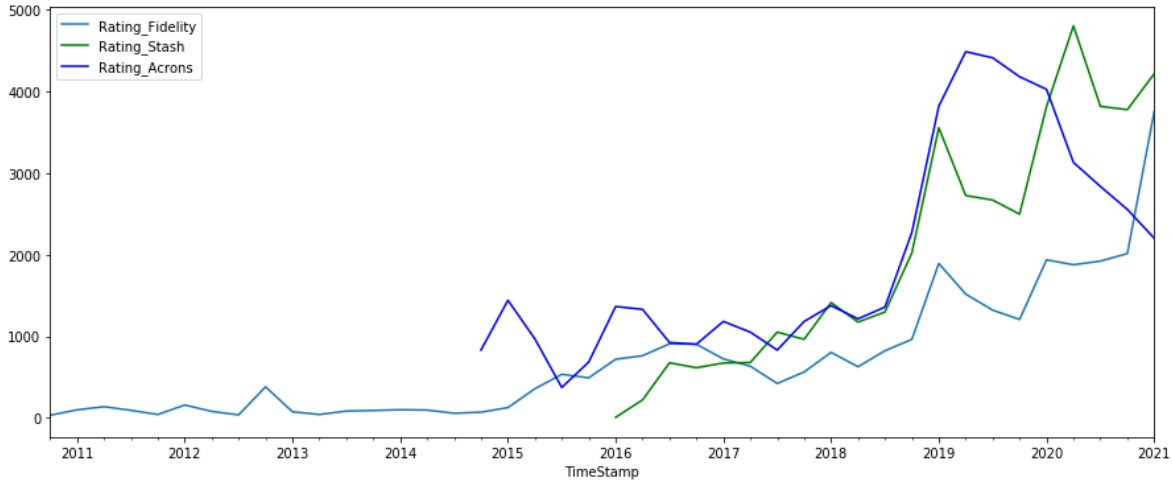
Figure 13: Finance Applications: Transaction, Rating Development



Source: Own creation

The rating development of the financial applications 'PayPal', 'CoinBase' and 'GooglePay' indicate stagnating to decreasing development in the form of average ratings in 2020. Only the application 'CashApp' indicates both increasing growth in the number of reviews and stable rating score development.

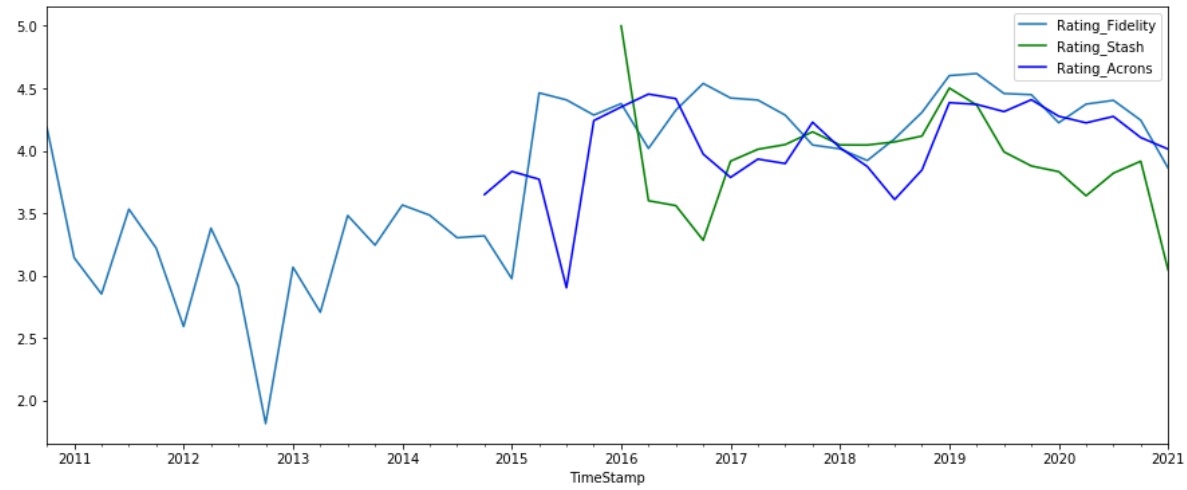
Figure 14: Finance Applications: Trading, Number of Reviews



Source: Own creation

In the segment financial applications for the type 'Trading', the applications 'Fidelity' and 'Stash' have their all-time peak number of reviews in 2020, with a positive growth trend towards the end of the year. The application 'Acrons' indicates a significantly decreasing growth rate throughout the whole of 2020.

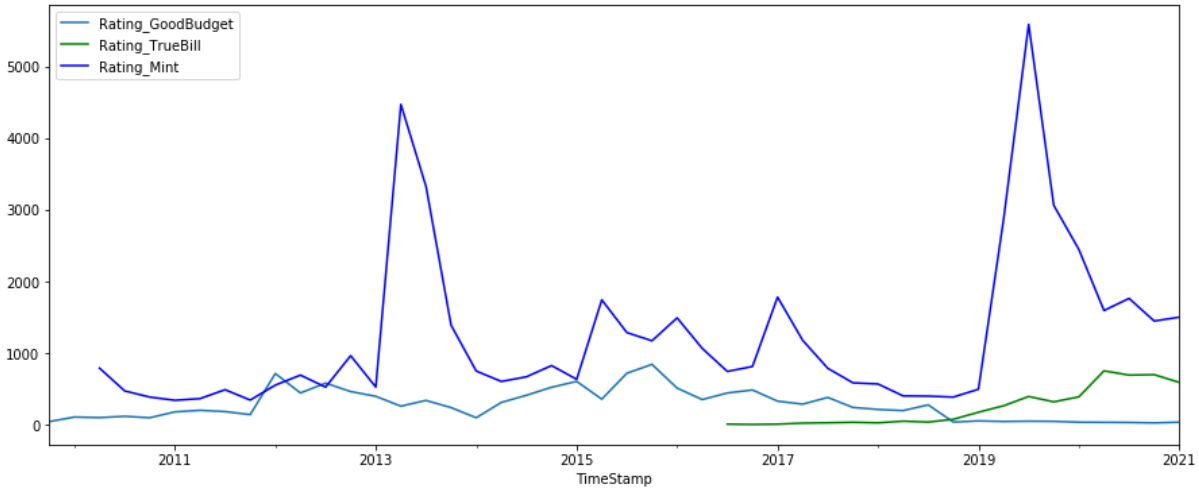
Figure 15: Finance Applications: Trading, Rating Development



Source: Own creation

The rating development of all applications indicates a stable rating average for the first three quarters of 2020, with a negative trend towards the end of the year.

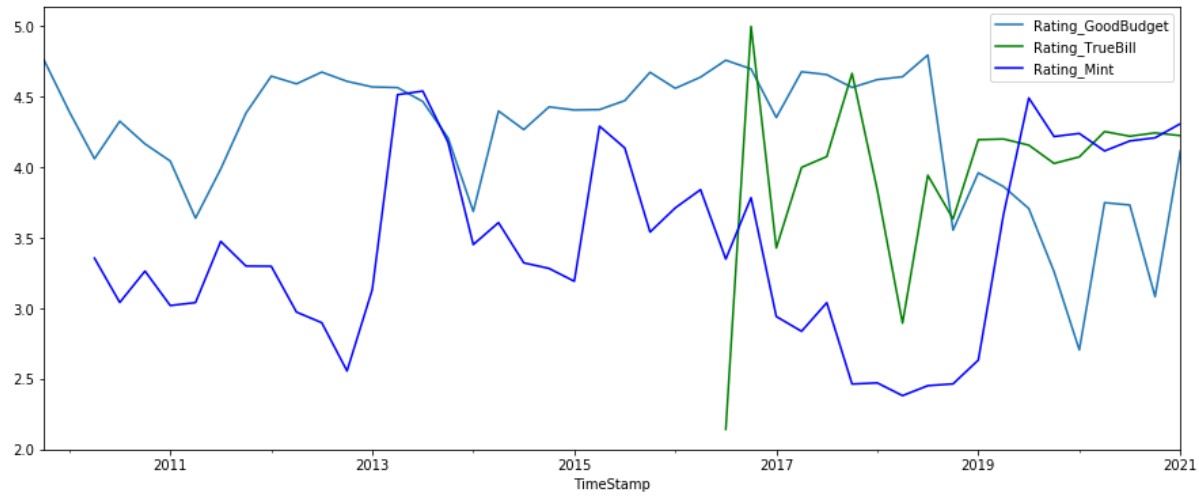
Figure 16: Finance Applications: Budget Planner, Number of Reviews



Source: Own creation

In the segment financial applications for the type 'Trading', the applications 'Mint' and 'GoodBudget' indicate a stable to slightly decreasing number of reviews throughout 2020. The growth rate of the application 'TrueBill' peaks in the year 2020.

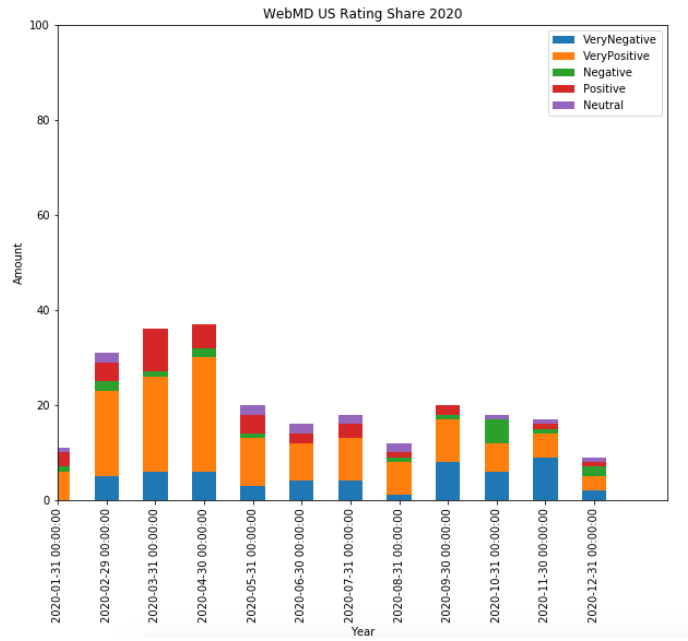
Figure 17: Finance Applications: Budget Planner, Rating Development



Source: Own creation

All applications indicate a stable to positive rating development towards the end of 2020. While 'Mint' and 'TrueBill' indicate a stabilized rating development, the application 'Goodbudget' indicates very volatile development throughout the year.

Figure 18:WebMD Review Distribution 2020

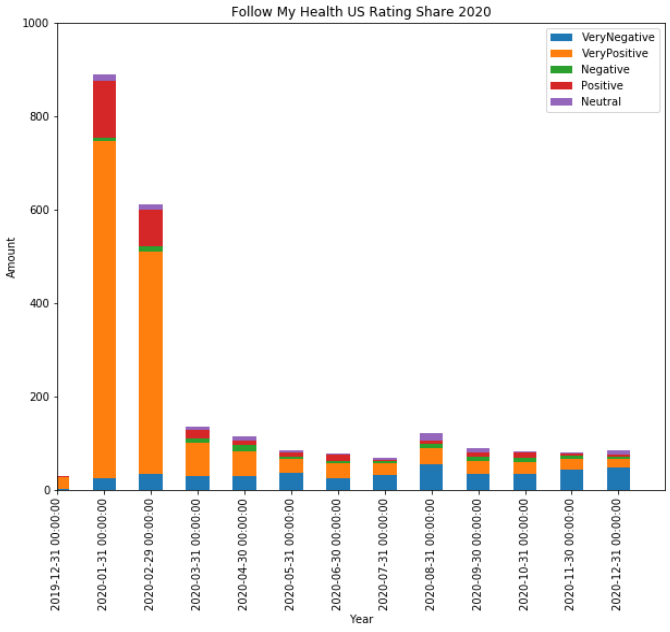


Source: Own creation

3.4 Identified Outliers

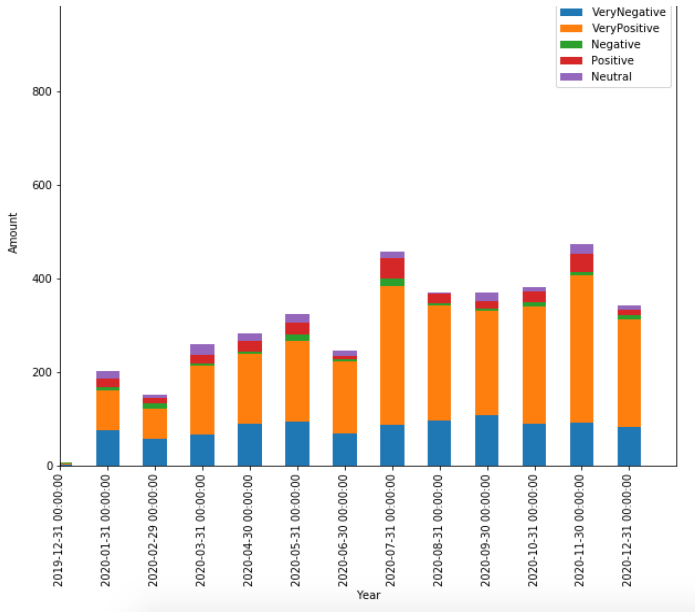
The identified outliers for each subgroup are as follows: Subgroup 'Medicine, Monitor' mHealth applications in the U.S.: No significant, comparable outliers were identified, although the application 'WebMD' indicates an unusual development at the end of 2020. Therefore more in-depth analyses were conducted. Figure 18 in fact indicates a higher negative number of reviews in the latter half of 2020. Still, the number is too small to draw valuable insights with a more in-depth analysis. Subgroup 'Telehealth' mHealth applications in the U.S.: The application 'FollowMyHealth' indicates a significant decrease in the number of reviews and average rating score for 2020, while 'MyChart' represents the opposite. Therefore a comparison over the duration of 2020 is valuable. For the telehealth application 'Follow My Health', a more in-depth analysis indicates that the beginning of year 2020 had a dominant number of positive. The drop in rating average presented in Figure 19 can be explained by the immense drop in the number of reviews, equaling the positive/ negative ratio. From September to the end of the year, the negative reviews represent a slightly higher number. Nevertheless, this should not be considered a major failure from the application provider side. Considering the small number of reviews and the negative/positive ratio, a more in-depth analysis did not seem valuable. The application 'MyChart' indicates a dominant numbers of positive reviews

Figure 19: Follow My Health Review Distribution 2020



Source: Own creation

Figure 20: MyChart Review Distribution



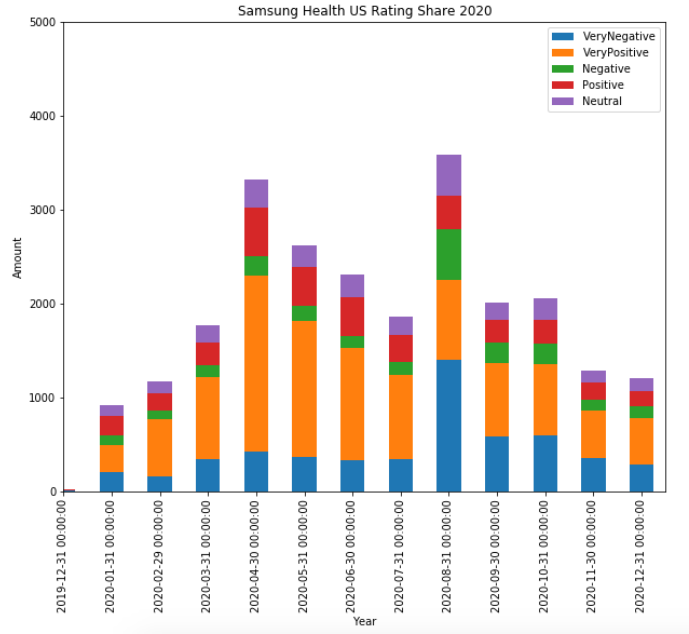
Source: Own creation

throughout the year. Although 2020 does not represent the highest average rating score of the application's lifespan, the respective high number of reviews represents a good example of what reviewers value. A more in-depth analysis for the whole year seems valuable.

The subgroup 'Health & Fitness' mHealth applications in the U.S.: The applications 'Huawei Health' and 'Samsung Health' both indicate significant growth in the number of reviews in 2020.

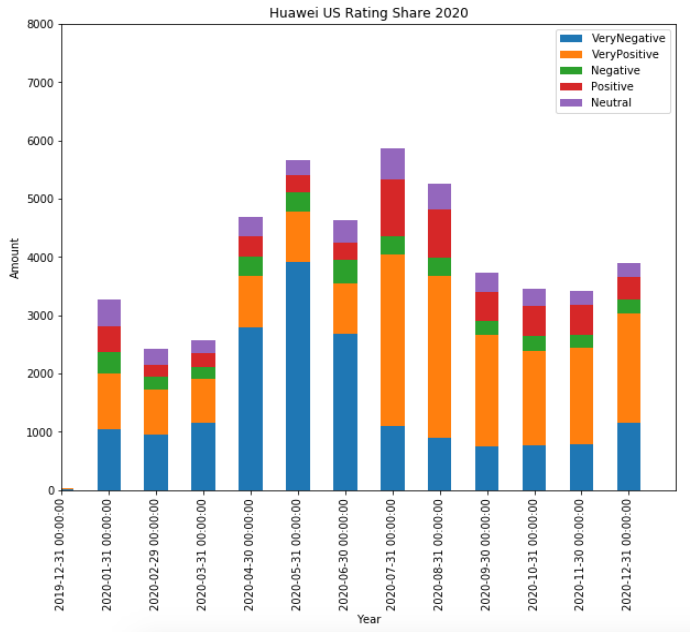
While 'Samsung Health' had the highest average rating score of its lifespan in the first quarter of 2020, 'Huawei Health' hit the lowest average rating score of its lifespan in the first quarter of 2020. 'Huawei Health' then recovered strongly for the rest of the year, reaching a new peak in the latter half of the year. 'Samsung Health' went in the opposite direction, indicating a strongly decreasing rating average for the rest of the year. The 'Samsung Health' application indicates a high number of positive reviews from February 2020 to the end of July 2020, followed by a drastic change at the beginning of September 2020. A more in-depth analysis of the reviews from February 2020 to July 2020 makes sense, in order to

Figure 21: Samsung Health Review Distribution 2020



Source: Own creation

Figure 22: Huawei Health Review Distribution 2020



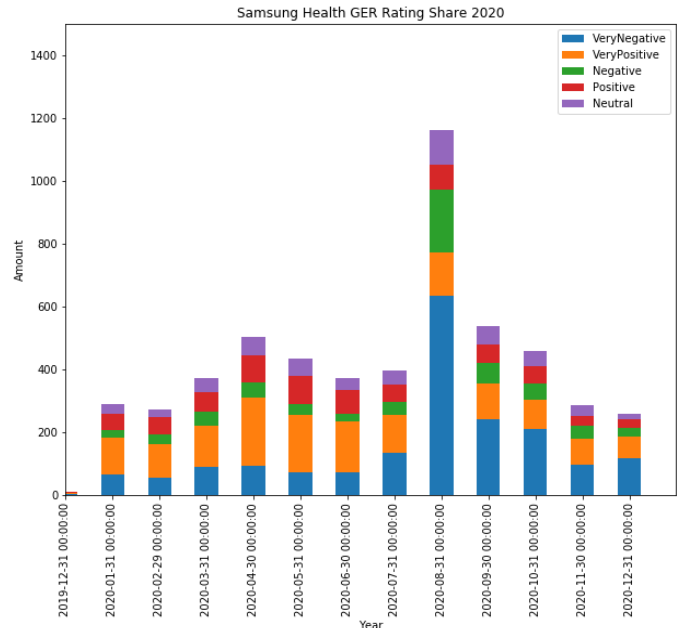
Source: Own creation

understand the reviewers' opinions regarding the application. In addition, the month of September 2020 should be further analysed, as it represents a major shift in the review score and a number of negative reviews, while representing the highest number of reviews in 2020.

The application 'Huawei Health' starts the first quarter of 2020 with a balanced positive/negative review ratio. From April to June 2020 there is a significant increase in negative ratings, while from July to the end of the year positive reviews strongly outweigh the negative. Further research for the reason for the drastic change in the middle of the year seems highly valuable.

Subgroup 'Medicine, Monitor' mHealth applications in Germany: No significant, comparable outliers were identified.

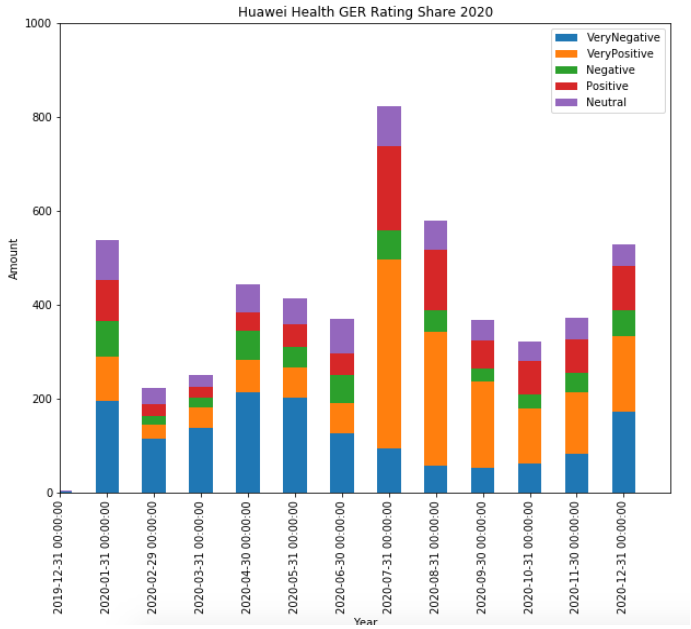
Figure 23: Samsung Health Germany Review Distribution 2020



Source: Own creation

Subgroup 'Health & Fitness' mHealth applications in Germany: The same strong outliers as in the U.S. market with 'Huawei Health' and 'Samsung Health'. As in the U.S. counterpart, the 'Samsung Health GER' application indicates a dominant number of positive reviews at the beginning of 2020, lasting until July. Contrary to the U.S. reviews, the trend of negative reviews continues until the end of 2020. A more in-depth analysis of the two trends seems

Figure 24: Huawei Health Review Germany Distribution 2020



Source: Own creation

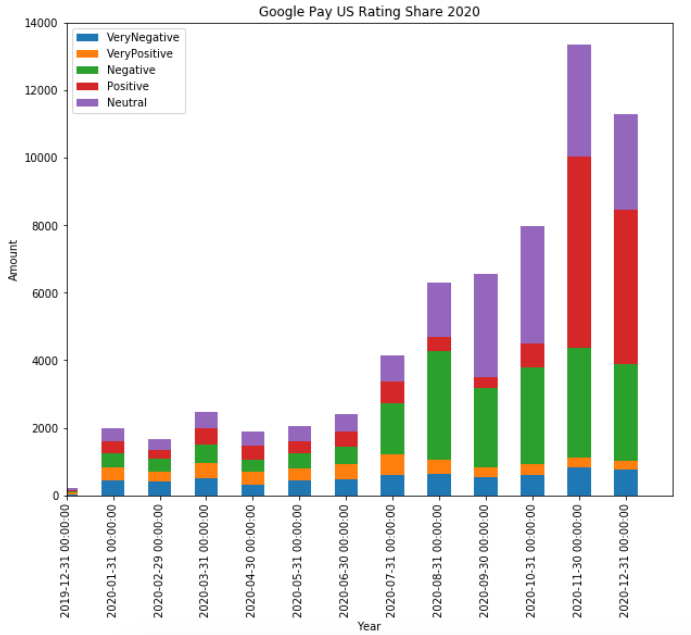
highly valuable, also in order to do a cross-national comparison. Again, like the U.S. counterpart, the German 'Huawei Health' application indicates a dominant number of negative reviews from the beginning of the year to August 2020, followed by a major shift to positive reviews. A more in-depth analysis again seems highly valuable.

Subgroup "'Trading"' of finance application: No significant, comparable outliers were identified.

Subgroup "'Budget Planner"' of finance application: No significant, comparable outliers were identified.

Subgroup 'Transactions' of finance application: The application 'Google Pay' indicates a significantly high growth rate during 2020. The average rating score indicates a significant drop in the third and fourth quarters, followed by a recovery phase to the end of the year. The application 'Google Pay' indicates a high number of negative reviews for the months August to October 2020. The last two months are marked by a significant increase in reviews, both negative and positive, where the positive reviews outweigh the negative ones. A more in-depth analysis of especially the last two seems highly valuable.

Figure 25: Google Pay Review Distribution 2020



Source: Own creation

3.5 Results: Research Question 2

Table 5: Word Frequency Analyses: Outliers

Application Subset	Unigram Positive Results	Bigram Positive Results	Trigram Positive Results	Unigram Negative Results	Bigram Negative Results	Trigram Negative Results
MyChart (2020)	**Easy**, helpful…	**Easy use**, , within minute…	**Easy use helpful**, app easy use…	**Time**, service…	**Speak doctor**, talk doctor…	**See doctor person**, , something go wrong …
Samsung Health (01-07, 2020)	**Help**, work…	**Keep track**, easy use…	**App easy use**, app help track…	**Update**, stress …	**Heart rate,** last update …	**Heart rate monitor**, since last update…
Samsung Health (08, 2020)	**Update,** feature…	**Weight management**, step count…	**Weight management feature**, weight management section…	**Weight,** update…	**Weight management,** management feature, late update…	**Weight management feature**, please bring back…
Huawei Health (01-07, 2020)	**Update**, work…	**Honor band**, watch face…	**Huawei mobile service**, app work well…	**Work,** update…	**Mobile service**, late version…	**Huawei mobile service**, last version Huawei…
Huawei Health (07-12, 2020)	**Use**, work…	**Easy use**, work well…	**App easy use**, app work well..…	**Update,** watch…	**Honor band,** update device …	**Update device list**, connect honor band…
Samsung Health GER (01-07, 2020)	**Zufrieden**, funtioniert.. .	**Zufrieden app**, macht spaß...	**App macht spaß**, , macht spaß motiviert...	Schritte, update...	**Letzen update**, funktioniert mehr...	**Seit letzten update**, letzten update funktioniert.
Samsung Health GER (08-12, 2020)	**Funktioniert**, spaß...	**Macht spaß,** app motiviert...	**Macht spaß motiviert,** nutze app seit ...	**Update**, Gewichtsverwaltung	**Letzen update**, Gewichtsverwaltung entfernt....	**Seit letzten update**, seit neuen update...
Huawei Health GER (01-07, 2020)	**Funktioniert**, update...	**Funktioniert einwandfrei**, huawei band...	**Huawei, mobile service**, huawei band pro... ..	**Update**, funktioniert ...	**Letzten update**, mobile service...	**Seit letzen update**, huawei mobile service,...

Huawei Health GER (08-12, 2020)	**Funktioniert**, zufrieden...	**Funktioniert gut**, funktioniert einwandfrei...	**Verbindung huawei watch**, app funktioniert gut...	**Funktioniert,** update...	**Letzen update**, uhr verbinden...	**Seit letzten update**, honor band koppeln....
Google Pay (07-12, 2020)	**Payment**, time …	**Easy use**, money transfer…	**Check account balance**, app money transfer…	**Payment,** time …	**Work properly**, bad experience …	**Take long time**, take much time…

Source: Own creation

MyChart (2020): The text analyses of the application 'MyChart' for 2020 indicate that the ease of use is the most dominant reason for positive reviews, followed by the quick response time and quality of service of doctors. Negative reviews focused on the actual output of the application, diagnoses and virtual meetings with doctors, issues which are hard to change from an application developer's perspective, as the actual therapy is most often not in the hands of the application. In addition, customer service was an often-cited reason for complaints.

Samsung Health (01–07, 2020): In the first half of 2020, 'Samsung Health' was marked by a dominant number of positive reviews. The usefulness and helpfulness of the application were dominant reasons for positive reviews. Reasons for complaints were mainly a new update, which had an impact on the functionality of the heart rate monitor (stress level), and the step count, resulting in a significantly increased number of negative reviews.

Samsung Health (08, 2020): For 'Samsung Health', the month September represents a major shift in the rating score average, with a decreasing number of positive reviews and a strongly increasing number of negative reviews. An in-depth analysis of the month indicates that the main reason for this change was a new update which removed the weight management option. This removal resulted in an uproar and a significant number of complaints. Even in the positive ratings, the removal of the weight management option was the dominant topic.

Huawei Health (01–07, 2020): In the first half of 2020, reviewers increasingly complained about an update which resulted in a server error when attempting to open the application. Also without the update, synchronization with other Huawei devices was not possible, which also resulted in displeasure. This is particularly problematic, as the synchronization of the application with other devices, such as the Honor Magic, was one of the main reasons for positive reviews.

Huawei Health (07–12, 2020): Although the application strongly recovers in the latter half of 2020, the synchronization of the mobile application with other devices, mainly the Honor Magic watch, is still the main reason for many complaints.

Samsung Health GER (01–07, 2020): Like their U.S. counterpart, the German 'Samsung Health' application receives many compliments for ease of use, combined with fun factors, driving the motivation for continuous usage. The new update was again the main reason for complaints, to be precise, errors with the step counter feature.

Samsung Health GER (08–12, 2020): Again, the main reason for the major shift in the average rating score was based, like with the U.S. version, on an update which removed the weight management option, resulting in a significant increase in complaints about this particular topic.

Huawei Health GER (01–07, 2020): As in the U.S., the first half of 2020 is highlighted by complaints about the latest update creating issues with synchronization, as well as the sleeping data transmission. Nevertheless, the overall functionality was the main reason for the positive reviews.

Huawei Health GER (08–12, 2020): In the latter half of 2020, 'Huawei Health' solved the main problems with the synchronization issues. The predominantly positive comments compliment the functionality and the simplicity of the interface. Nevertheless, synchronization still remains an issue for many customers, as it remains the main topic of negative reviews.

Google Pay (07–12, 2020): The latter half of 2020 started with a predominant number of negative reviews, indicating a strong shift in the last two months of the year. The main reason for the negative reviews was performance-related issues, especially technical issues during transactions.

3.6 Summary Text Analyses all Subgroups

Table 6: Word Frequency Analyses: mHealth Applications U.S.

Application mHealth U.S. Subgroups	**Unigram Positive Results**	**Bigram Positive Results**	**Trigram Positive Results**	**Unigram Negative Results**	**Bigram Negative Results**	**Trigram Negative Results**
Medicine Monitor	**Helpful**, help…	**Easy use**, helpful app…	**App easy use**, help keep track…	**Work,** sign update…	**Force close**, download app…	**Past first screen**, every time try…
Telehealth	**Easy**, helpful,…	**Easy use**, keep track …	**App easy use**, great way keep,…	**Time**, work, log, e-mail,…	**Download app**, customer service,…	**Something go wrong**, every time tr …
Health & Fitness	**Work** help …	**Easy use**, keep track …	**Help keep track**, keep track med …	**Update**, work...	**Count step**, heart rate…	**Since last update**, weight management featuree…

Source: Own creation

Medicine, Monitor mHealth U.S.

The text analysis for this subgroup indicates that customers value the ease of use and usefulness the most, as these were the major reasons for positive and negative reviews. In particular, the sign-in and account creation process were accompanied by numerous bugs and errors, resulting in dissatisfaction.

Telehealth mHealth U.S.

The subgroup 'Telehealth U.S.' indicates a high satisfaction level, based on ease of use and helpfulness. Topics such as questions, customer service and the onboarding process were the main reasons for dissatisfaction.

Health & Fitness mHealth U.S.

The subgroup 'Health & Fitness U.S.' indicates a high satisfaction level, based on the functionality, ease of use and usefulness. Erroneous update transitions, where features are removed or resulting in errors, are the main reasons for dissatisfaction.

Table 7: Word Frequency Analyses: mHealth Applications Germany

Application mHealth German Subgroups	**Unigram Positive Results**	**Bigram Positive Results**	**Trigram Positive Results**	**Unigram Negative Results**	**Bigram Negative Results**	**Trigram Negative Results**
Medicine Monitor	**Einfach**, genau...	**Einfach bedienen**, leicht bedienen...	**Finde app einfach**, app einfach bedienen...	**Daten**, funktioniert…	**Letzten update**, seit letzten...	**Seit letzten update**, seit neuen update...,
Health & Fitness	**Funktioniert**, einfach...	**Macht spaß**, schöne app,...	**Seit letzten update**, app macht spaß,...	**Update**, funktioniert...	**Letzten update**, seit letzen...	**Seit letzten update**, letzten update funktioniert...

Source: Own creation

Medicine, Monitor mHealth Germany

Again, the ease of use, usefulness and well-arranged interfaces were the dominant reasons for positive reviews. Updates with associated errors during the sign-in process, as well as data protection deficiencies, were the main reasons for negative reviews.

Health & Fitness mHealth Germany

The ease of use, usefulness, functionality and well-arranged interfaces, combined with the fun factor, resulted in higher motivation and a high number of positive reviews. Updates, in combination with feature removals and errors, were the main reasons for negative reviews.

Table 8: Word Frequency Analyses: Finance Applications

Application mHealth U.S. Subgroups	Unigram Positive Results	Bigram Positive Results	Trigram Positive Results	Unigram Negative Results	Bigram Negative Results	Trigram Negative Results
Transactions	**Easy**, work…	**Easy use**, quick easy …	**App easy use**, great app easy…	**Account**, customer…	**Customer service**, customer support…	**Link bank account**, add bank account…
Trading	**Easy**, account…	**Easy use**, user friendly …	**App easy use**, great app easy…	**Account**, customer…	**Customer service**, customer support…	**Get money back**, call customer service …
Budget Planner	**Account**, easy…	**Easy use**, one place…	**Help keep track**, keep track finance…	**Account**, updated…	**Bank account**, update account …	**Link bank account**, connect bank account …

Source: Own creation

Transactions Finance

The ease of use, the functionality in the form of speed and the quick response times were the dominant factors for positive reviews. Account and transaction issues, combined with bad customer service, were the main reasons for negative reviews.

Trading Finance

Again the ease of use was the most dominant factor of positive reviews. Bank account issues, in combination with poor customer support, presented the main reasons for negative reviews.

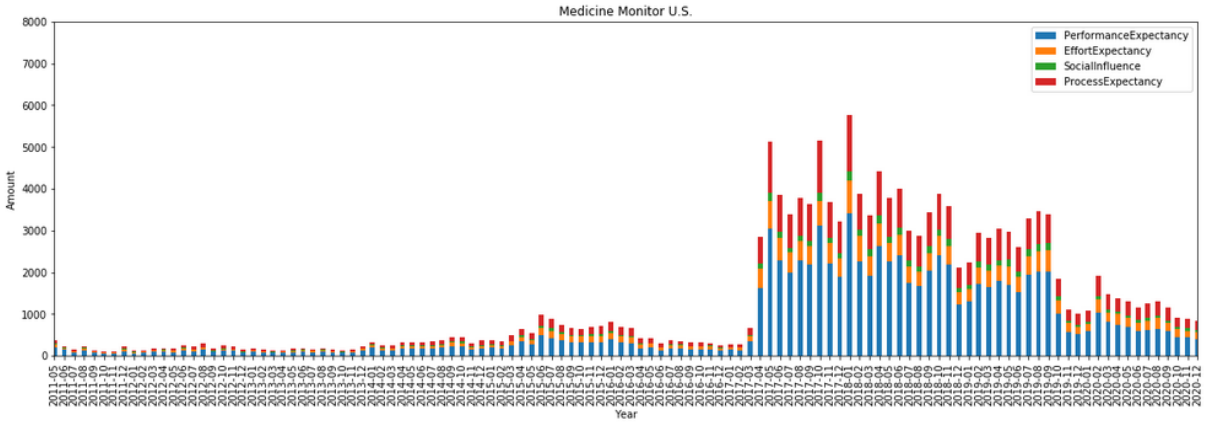
Budget Planner Finance

The helpfulness and ease of use were the main reasons for positive reviews. New updates, in combination with errors during the account creation and maintenance process, were the dominate reasons for negative reviews.

3.7 Results: Research Question 3

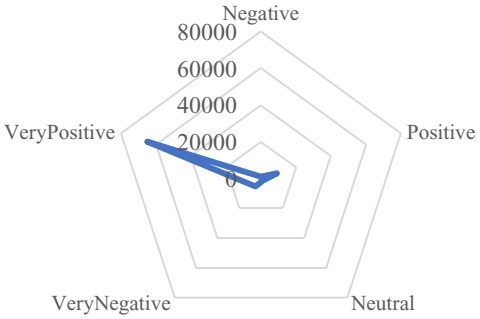
Medicine Monitor U.S.

Figure 26: Medicine Monitor, U.S., Key Features Distribution



Source: Own creation

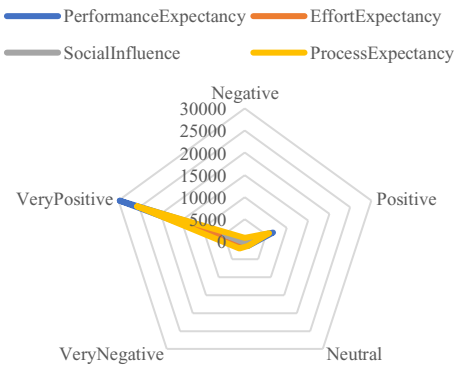
Figure 27: Rating Distribution over the Whole Maturity, Medicine Monitor U.S.



Source: Own creation

The strongest key feature for the 'Medicine Monitor' subgroup in the United States is the 'performance expectancy', indicating that the actual output/performance of each application is the focus for most reviewers. This implies that categories such as functionality, satisfaction level, outcome/results of the respective services, usefulness and so on are the most dominant reasons for reviews/ comments in the subgroup 'Medicine Monitor', followed by the 'process expectancy' as the second strongest key feature. One should consider that categories from 'process expectancy', such as malfunctions, errors and/or new dysfunctional features can result in a direct impact on the 'effort and performance expectancy'. For example, a new update increases the complexity of the application's interface, resulting in increased negative perceptions on the part of reviewers, generating complaints about the functionality and ease of use, and therefore impacting the category 'effort and performance expectancy'.

Figure 28: Key Feature Distribution over the Whole Maturity, Medicine Monitor U.S.



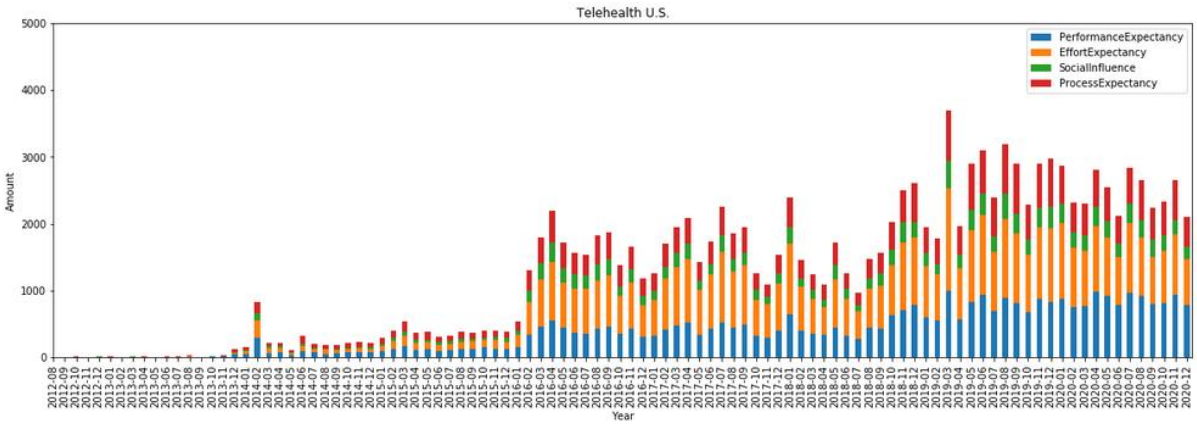
Source: Own creation

Nevertheless, the structure of the analysed reviews most often indicates a strong distinctiveness of each key feature. Application developers in this particular subgroup should focus their efforts on ‘performance and process expectancies’; especially functionality in combination with high-quality outputs.

When splitting the data sets into positive and negative reviews, the key features ‘performance and process expectancy’ were the most dominant reasons for negative reviews, mainly based on functional errors during the sign-in process. However, the overall perception is highly positive, with a high satisfaction level in regards to ‘performance and process expectancy’. The simplicity and functionality of the offered services were regarded highly. Maintaining the quality while focusing on customer acquisition should be the strategic focus. Cooperation with policymakers and insurances which assess quality output, functionality, safety, data security, data protection and positive effects on care would result in an official quality stamp, attracting new customers while requiring only low investments based on the overall highly regarded output quality.

Telehealth U.S.

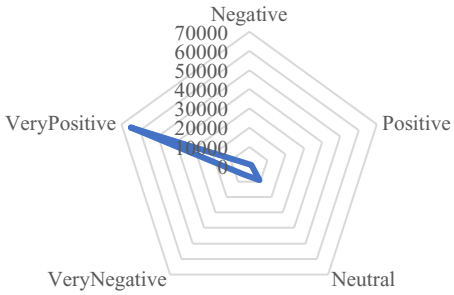
Figure 29: Telehealth, U.S., Key Features Distribution



Source: Own creation 1

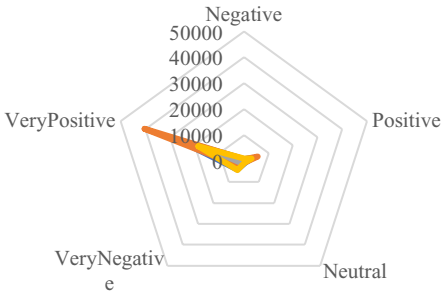
The strongest key feature in the 'Telehealth' subgroup is 'effort expectancy': Given the nature of telehealth applications, a seamless, easy and quick process with regard to its services, in the form of appointments, videocalls and therapies, is necessary to satisfy customers. The key features 'process and performance expectancy' are ranked second and third from the reviewers' perspective, indicating an important role for both the output of the application as well as seamless processes. In addition, the key feature 'social influence' has a high number of occurrences when compared to other subgroups. This indicates that there is a high number of word-of-mouth recommendations in this particular subgroup. Following the insights of the key feature analyses, application developers of this particular subgroup should continue their focus on 'effort and process expectancies'. Seamless processes for video calls, appointments, connectivity and customer services, resulting in a flawless service

Figure 30: Rating Distribution for the Whole Maturity Telehealth U.S.



Source: Own creation

Figure 31: Key Feature Distribution over the Whole Maturity, Telehealth U.S.



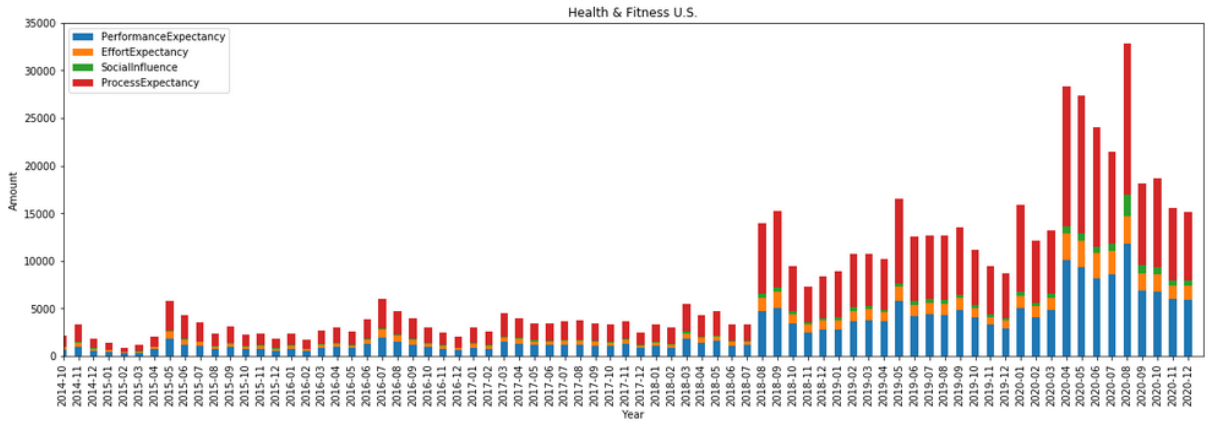
Source: Own creation

offering, will generate an increasing number of satisfied customers.

When splitting the data set into positive and negative, the feature ‘effort expectancy’ was the dominant reason for positive comments, while the features ‘performance and process expectancy’ were the main reasons for negative reviews. In addition, the factor ‘social influence’ has a higher occurrence in good reviews, indicating that a satisfying experience has a postive impact on recommendations. Given the unique telehealth situation in the U.S., allowing cross-state services, preserving these regulatory measurements, should be in the interests of insurances, healthcare systems and mHealth application developers.

Health & Fitness U.S.

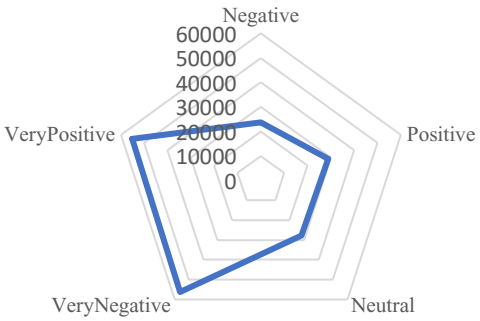
Figure 32: Health & Fitness, U.S., Key Features Distribution



Source: Own creation 2

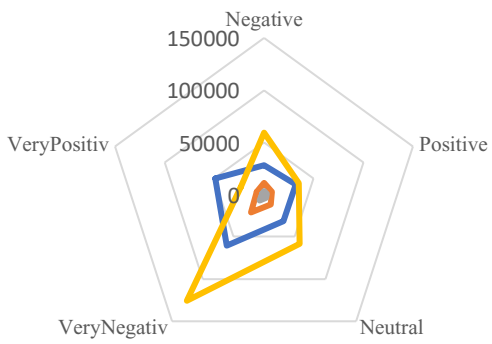
The ‘Health & Fitness’ subgroup is dominated by the ‘process expectancy’ key feature. As the previous analysis indicated, new updates, changing the interface, reducing features and resultant errors where the reasons for a major uproar in the reviews. The radar diagram indicates that the complaints were far more extensive in length, explaining the gap between ‘VeryPositive’ and ‘VeryNegative’ perceptions. Especially in 2020, newly implemented updates were reasons for a significant increase in negative reviews. Before developing and implementing new updates which include new features and/or drop old features, a rigorous failure frequency

Figure 33: Rating Distribution for the Whole Maturity Health & Fitness U.S.



Source: Own creation

Figure 34: Key Feature Distribution for the Whole Maturity Health & Fitness U.S.

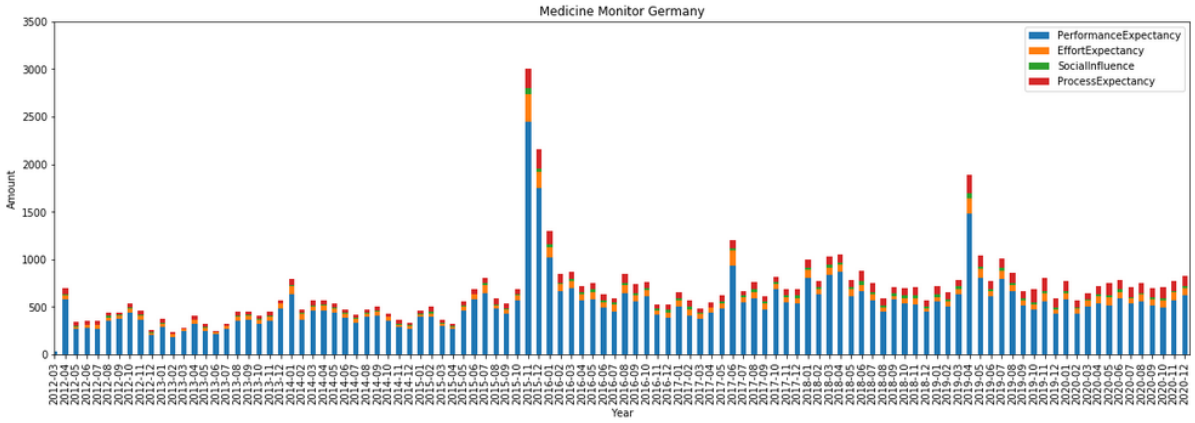


Source: Own creation

analysis, as well as extensive market research should be conducted, in order to generate a seamless change process for each new update. Given the nature of 'Health & Fitness' applications, being often interlinked with external devices, a flawless connection process should be an additional focus.

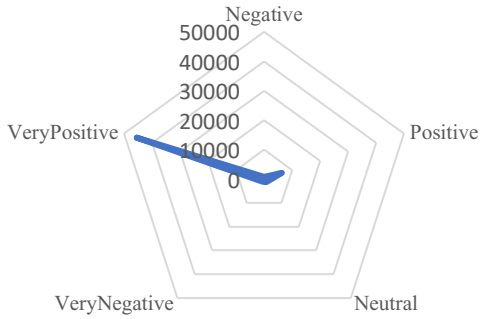
Medicine Monitor Germany

Figure 35: Medicine Monitor, Germany, Key Features Distribution



Source: Own creation

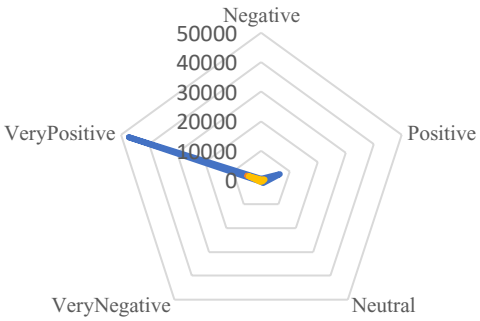
Figure 36: Rating Distribution over the Whole Maturity, Medicine Monitor Germany



Source: Own creation

Reviews of the German 'Medicine Monitor' application indicate a clear tendency towards 'performance expectancy' being the most dominant factor for writing a review, followed by 'effort and process expectancy' sharing the second rank. Based on these insights, a general high satisfaction level can be concluded. A clear distinction of key feature preference based on negative or positive reviews can be identified. Comments about 'process

Figure 37: Key Feature Distribution over the Whole Maturity, Medicine Monitor Germany

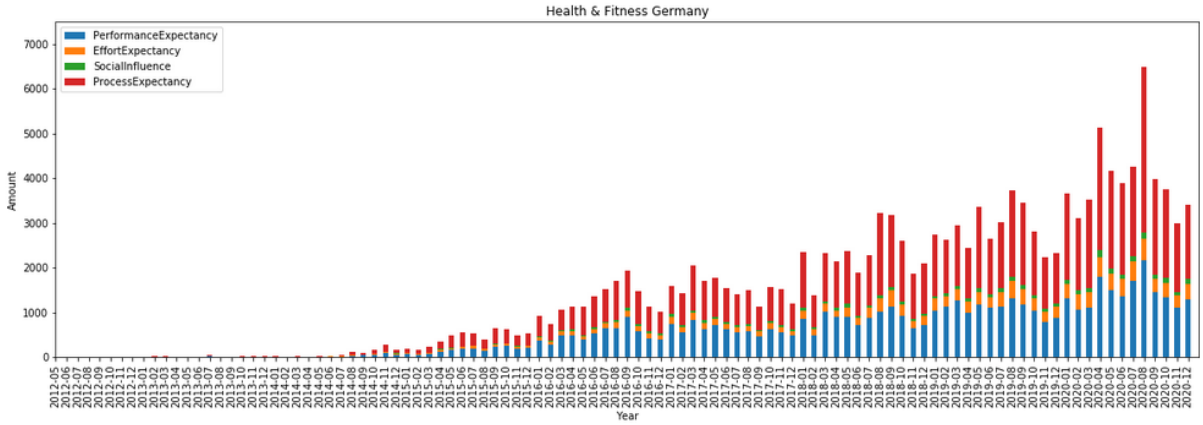


Source: Own creation

expectancy' are the reason for negative reviews, while for positive reviews 'performance expectancy' is the strongest factor. Strategic measurements to avoid 'process expectancy' failures can be a crucial factor for the average rating score development. Although the overall positive perception outweighed the process-related issues, the offerings can be optimized by decreasing technical failures, while maintaining the status quo. Applying for the German BfArM assessment could be a fitting solution, without making higher investments.

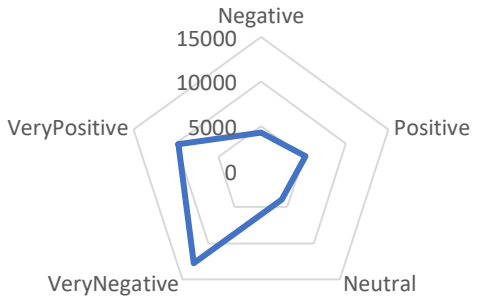
Health & Fitness Germany

Figure 38: Health & Fitness, Germany, Key Features Distribution



Source: Own creation

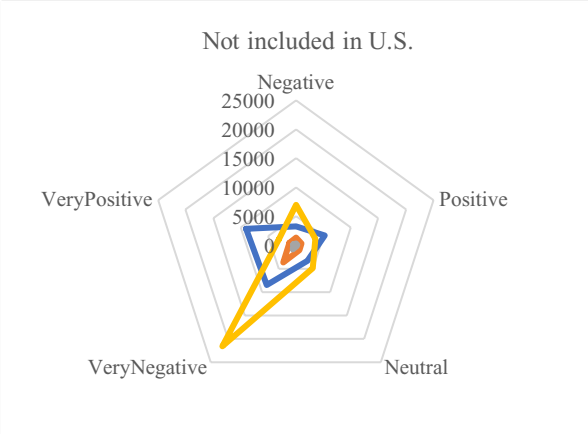
Figure 39: Rating Distribution over the Whole Maturity, Health & Fitness Germany



Source: Own creation

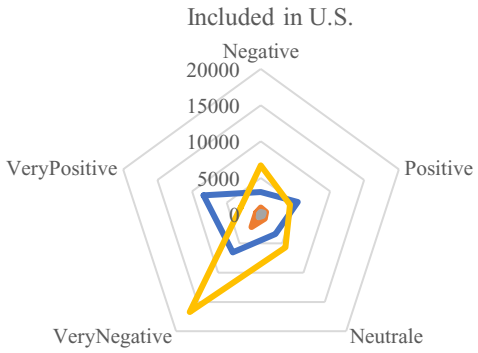
Like their U.S. counterpart, the 'Health & Fitness' applications also indicated a high level of dissatisfaction based on new updates, and removing or changing old key features of the respective application, resulting in errors and malfunctions. This occurrence applies for all applications, the ones used for the U.S. analysis as well as the additional applications. Accordingly, in the cross-country comparison,

Figure 40: Key Feature Distribution over the Whole Maturity, Health & Fitness Germany



Source: Own creation

Figure 41: Key Feature Distribution over the Whole Maturity, Health & Fitness Germany

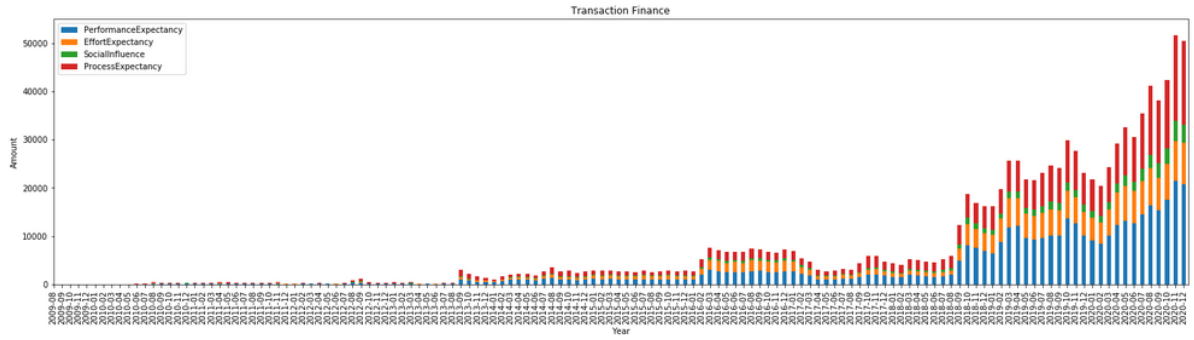


Source: Own creation

a clear distinction between the United States and Germany in the form of review preferences, could not be identified. Only the distribution of the key feature 'process expectancy' indicates small differences in the number of positive and negative reviews. Contrary to their U.S. counterparts, the German reviewers value the 'performance expectancy' slightly more. Also, this subgroup should focus their strategic measurements on rigorous planning of new updates, while taking some degree of noise in the change process into account. As the degree of technical complexity has a direct impact on reviewers' perceptions, transforming this vulnerability into a unique selling point would attract more customers, while differentiating themselves from competitors.

Transaction

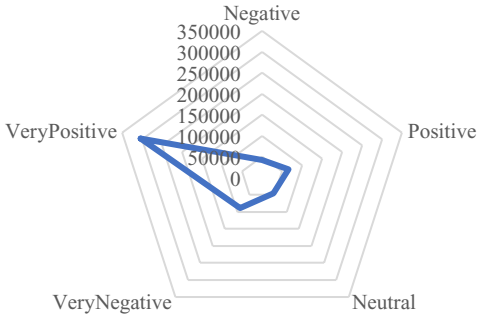
Figure 42: Transaction, Finance, Key Features Distribution



Source: Own creation

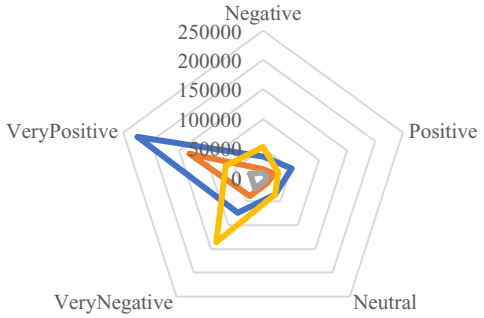
The key feature analysis for the subgroup 'Transactions' indicates a clear focus on 'performance and process expectancy' in the form of customers' reviews. The word frequency analyses indicated that malfunctions in the areas of transactions, bank account, and sign-in process, combined with a bad customer service experience, result in an increasing number of negative reviews. When dividing the data set into positive and negative, 'performance and effort expectancy' can clearly be seen as the features with the highest positive occurrence, indicating that the simplicity of the applications presents a high value from the customer perspective. Malfunctions and errors during the transaction process can have a strong impact on users' lives, resulting in negative reviews. Minimizing these malfunctions while increasing efforts for successful customer relationship management, thriving for excellence in both areas, would keep the main reasons for negative reviews to a minimum.

Figure 43: Rating Distribution over the Whole Maturity, Transaction



Source: Own creation

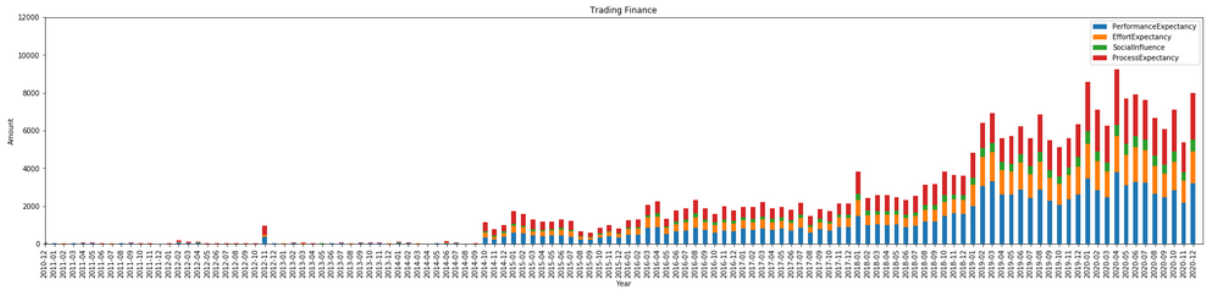
Figure 44: Key Feature Distribution over the Hole Maturity, Transaction



Source: Own creation

Trading

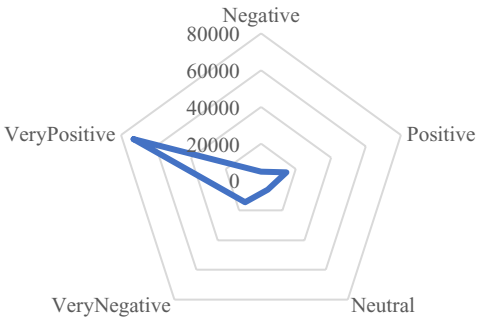
Figure 45: Trading, Finance Key Features Distribution



Source: Own creation

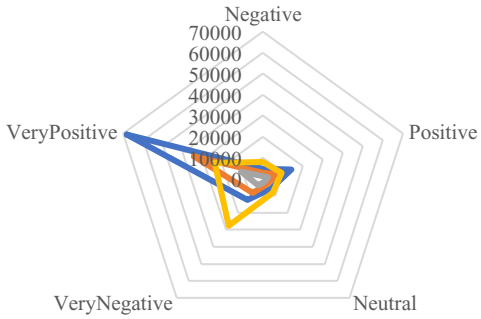
For the data set of the subgroup 'Trading', with only negative reviews the key feature 'process expectancy' had the most occurrences, while the data set with only positive reviews is dominated by the 'performance and effort expectancy'. Especially account issues, paired with unsatisfactory customer service, were the main reasons for negative reviews. This indicates that the main focus of the user lies with the ease of use, in particular during the sign-in process and the account maintenance/ closure. Simplifying the on-boarding and off-boarding process, while focusing on excellent service and customer support, can positively influence the perception of users. In addition, the feature 'social influence' had a comparably strong occurrence. Word-of-mouth recommendations could represent a strong marketing tool and should be considered.

Figure 46: Rating Distribution over the Whole Maturity, Trading



Source: Own creation

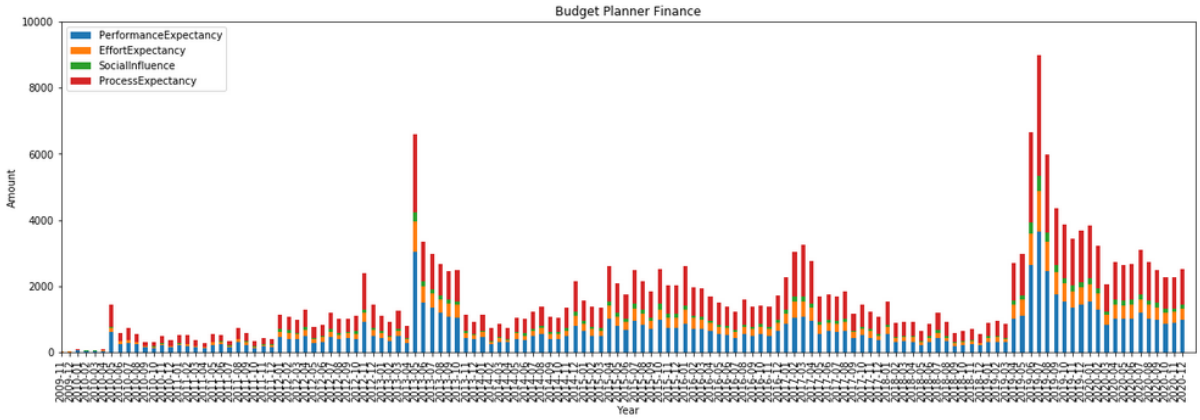
Figure 47: Key Feature Distribution over the Whole Maturity, Trading



Source: Own creation

Budget Planner

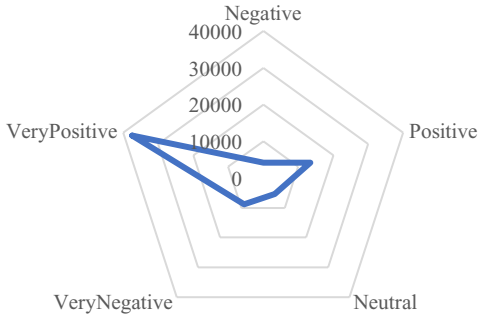
Figure 48: Budget Planner, Finance, Key Features Distribution



Source: Own creation

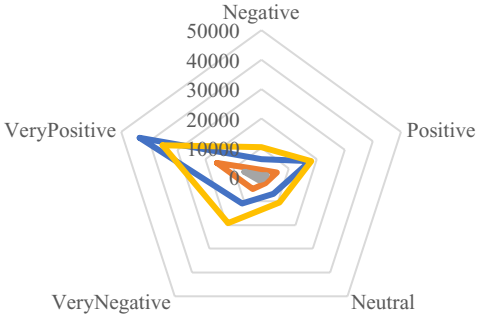
For the applications in the subgroup 'Budget Planner', the main focus of the reviews lies with the 'performance and process expectancy': especially new updates, impacting the functionality of the application, were reasons for negative reviews. To be more specific, new updates negatively impacting the functionality of one's personal account and the connectivity to the respective bank account were main reasons for complaints. With the positive comments, the focus was on the perceived usefulness and technical process of the application. Again, focusing on seamless transitions when introducing new updates should be of high importance.

Figure 49: Rating Distribution over the Whole Maturity



Source: Own creation

Figure 50: Key Feature Distribution over the Hole Maturity, Budget Planner



Source: Own creation

4 Discussion

4.1 Addressing the Research Questions

In this study, the evolution of mHealth applications, based on the number of reviews and rating scores in the year 2020, was investigated. The findings presented numerous insights into the review evolution during the pandemic year, as well as customer preferences for each analysed subgroup.

The results for the **first research question** indicated, for the cross-country comparison, no significant differences in the evolution of reviews or rating score for the subgroup 'Medicine Monitor' and the subgroup 'Health & Fitness'. Only the general number of reviews in Germany was significantly lower compared to the U.S., at some point even up to seven times lower. This difference cannot be explained by the difference in market size.

Considering the cross-industry analyses, the only subgroup which could have been impacted by the pandemic are the 'Trading' applications. They are comparable with the 'Telehealth' applications in the U.S., indicating similar trends in regards to number of reviews and rating score. In general, the insights of the number of reviews and the average rating score analyses were not able to clearly identify a trend towards a positive perception of mHealth applications. The results of the number of reviews development presented are contrary to many existing industry reports from McKinsey and Roland Berger, indicating a strong increase in and positive trend for mHealth adoption(McKinsey, 2020; Roland Berger 2020). Although the existing literature states that the overall download rate of mHealth applications increased significantly, the download rate and number of reviews do not necessarily have to correlate. One has to consider that the analysed applications had an above-average rating score and were generally well received by their users. Responsible for this was the decision-making process for choosing the applications, as the average rating score has a direct impact on the number of users and reviews. Further research should focus on the application retention rate, in order to identify which customers use the applications over a duration of 30 to 90 days. Based on the frequency of usage, customers and customer reviews need to be clustered. This will allow application developers to identify which customers and customer reviews should be used for further analyses, as they represent more accurate data.

The **second research question** identified outliers and gave insights into the reasons for these particular developments. The in-depth analyses of the outliers indicated that technical problems are the most dominant reasons for complaints and therefore responsible for a volatile development in average rating scores. The mentioned technical problems are most often based on new updates, which resulted in errors, malfunctions or the removal of specific features. The specific errors and removals differ from application to application: some of them relate to synchronization and connection errors, while others concern the sign-in and account creation processes. Dissatisfaction with these issues is often paired with an unsatisfactory customer service that is not able to solve the problems. The results indicated that applications with a higher degree of technical complexity (Health & Fitness and Transaction applications) are more vulnerable to malfunctions and errors. Accordingly, a higher negative review ratio is the result. Applications with low technical complexity were highly regarded, based on their performance output and simplicity. The reason for this could be that customer expectations were lower and easier to satisfy. Applications with a high degree of technological complexity have higher customer expectations. Technical features based on connectivity and synchronization with wearables are expected to work without flaws, explaining why customers value performance and process outcomes the most, as revealed by either positive or negative reviews.

The **third research question** gave insights into customer preferences, while presenting an easy-to-adapt NLP tool for further, more customized research. Although the UTAUT framework has not been used in the context of NLP analysis to date, the results were quite similar to the existing literature, which comments that determinants such as functionality, simplicity and required effort are of utmost importance from the customer's perspective (Shaoolian, 2017). Functional and technological issues, paired with high complexity and unsatisfactory services / customer services, were the main reasons for complaints. Current and future application developers should strongly consider these factors in order to maintain high quality and a high satisfaction level. The used natural language processing tools enabled the analysis of over a million customer preferences. Given the always growing amount of data, data analyses using online data will become increasingly important for future research.

The in-depth word-frequency analyses demonstrated their value by determining the reasons for either high satisfaction or dissatisfaction in a comprehensive way. For researchers, these tools represent an opportunity to fully understand mHealth customers' preferences, without having

access to company data. Future academic research should utilize this method, in order to improve the data amount and quality without having access to industry-specific data.

4.2 Implications and Recommendations

Given the increased acceptance of digital solutions from the side of patients and healthcare providers, the unique situation of the pandemic could be further utilized by application developers, policymakers, regulators and insurances. Even closer cooperation could enforce mHealth applications as a permanent, interchangeable component of future healthcare systems, generating benefits for all parties. In addition, the NLP tools used in this study demonstrated how valuable insights can be gained. The ubiquitous and always increasing amount of data produced by mHealth applications will help future stakeholders to fully understand customer preferences and act accordingly.

4.2.1 Companies and Managers

This research study indicated that the overall composition of mHealth applications, whereby focusing on the expected outcome, functionality, technical complexity and ease of use, should be used to classify the applications and adopt the most fitting strategy. It is important to comprehensively view every aspect and cluster these according to relevance from the client's perspective. For example, Health & Fitness applications have a higher technical complexity than Telehealth applications. This is based on numerous features that require more technically complex processes, which make the application more vulnerable to errors and malfunctions. Thus performance and outcome satisfactions are strongly interlinked with functionality. Telehealth applications, by comparison, have a low degree of technical complexity. Video calls in general require less technically complex processes, with a low failure rate, which the functionality of the application is most often not responsible for. Thus seamless functionality should be a given for these kinds of applications and the focus of customers' shifts to the actual output quality of the offered services. The used NLP tools only need small adjustments in order

to be feasible for other types of application, opening opportunities for other application providers or B2C services with online reviews.

4.2.2 States and Regulators

Many countries and their regulators already understand the great value of digital health solutions in the form of mHealth applications. Governments and insurances have increased their efforts and communications with mHealth application developers by including them in insurance plans, making major investments and creating regulatory measurements in order to assess mHealth applications. This study supports policymakers and regulators to classify applications and improve their assessment process. Considering the high average rating score of the analysed applications, closer cooperation between policymakers/ regulators and application developers is advisable in order to attract more customers, relieve stress and costs for healthcare systems and improve the overall quality of health for citizens.

4.2.3 Researchers and Scholars

The presented NLP tools can assist future researchers and scholars as a foundation for text analysis in nearly every B2C sector. As already mentioned, the necessary background knowledge was kept to a minimum, allowing researchers with no history in this field to use the taxonomy according to their preferences and chosen topic. The results of the study gave a comprehensive overview of the development of mHealth applications, indicating the high value and vast amount of online data. The resulting insights can be used for further research focusing on the micro level of each application. Customized text analyses for each application would go beyond the scope of this thesis, but could provide very precise information about customer behavior. Adding additional information to the data set would allow external observers limited insights into normally internally accessible industry information. Rightfully used, analyses of unprecedented heights can be down by private researchers.

4.3 Limitations and Future Research

The limitations in answering RQ1 are based on the limited access to data. The number of reviews and ratings given with each review represent only a fraction of the total download rate and given ratings. In order to fully answer the question, one needs to have access to the monthly download and dropout rates, as well as all the ratings given in general, in order to identify the

application retention rate. With this data with higher accuracy, quality and value can be used for further research. Regarding the key feature analyses, a more distinctive model is advisable; although the UTAUT model has already been used in existing research, filling individual components with a dictionary may result in overlaps. For example, the term ‘simple’ can be used for either ‘effort, performance and/or process expectancy’. Creating customized features and an aligned dictionary was beyond the scope of this work; however, a taxonomy for future, in-depth analyses was presented. Future work can now focus on individual applications, adapting the insights of the word frequency analyses to individual features. Therefore more detailed opinion-mining tools from NLP, such as a fine-grained analysis based on unsupervised or semi-supervised approaches, such as the latent Dirichlet allocation (LDA) and its variants, can be used. In addition, sentiment analyses, emotion detection, aspect-based sentiment analyses and multilingual sentiment analyses can further enhance more in-depth opinion mining.

5 Conclusion

The findings if this study suggest that the perception of mHealth applications vary strongly according to their purpose, segment and composition, whereby the influence of the pandemic was not clearly decteable. The applications are valuable and held in high regard if they are simple and intuitive to use. Especially for applications where higher technical complexity is a core component of certain services and features, the flawless execution of these is of vital importance.

A comprehensive analysis of the applications’ capabilities and the respective customer preferences is necessary to reach the highest satisfaction level. Further investigation into mHealth application reviews will be worthwhile, given the vast amount of information accessible without requiring industry-specific access.

mHealth providers should take the initiative and seek to intensify their cooperation with insurances and policymakers, to comprehensively integrate more mHealth services into healthcare systems. At the same time, they should continue to answer customers’ needs and wants by reducing costs and improving output quality.

Despite the increasing interest in the workings and improvement of NLP tools for online review analyses, there are still gaps in the literature, in regards of a simple to use taxonomy. One aim

of this study was to address the usefulness of NLP tools for private usage. In the future, studies should focus on individual applications, using more customized NLP tools and external company information in order to identify application-specific trends and customer preferences/ behavior. Following a top-down approach, future researchers could analyse whether the application-specific customer perspectives and trends apply to peer applications.

Appendix Literature

Akter, S., & Ray, P. (2010). mHealth - an Ultimate Platform to Serve the Unserved. *Yearbook of medical informatics*, 94–100. DOI: 10.1055/s-0038-1638697

Alaiad, A., Alsharo, M., & Alnsour, Y. (2019). The Determinants of M-Health Adoption in Developing Countries: An Empirical Investigation. *Applied clinical informatics*, 10(5), 820–840. https://doi.org/10.1055/s-0039-1697906

COVIDSurg Collaborative (2020). Elective surgery cancellations due to the COVID-19 pandemic: global predictive modelling to inform surgical recovery plans. *The British journal of surgery*, 107(11), 1440–1449. https://doi.org/10.1002/bjs.11746

Free, C., Phillips, G., Felix, L., Galli, L., Patel, V., & Edwards, P. (2010). The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. *BMC research notes*, 3, 250. https://doi.org/10.1186/1756-0500-3-250

Gerke, S., Stern, A.D. & Minssen, T., Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. *npj Digit. Med*. 3, 94 (2020). https://doi.org/10.1038/s41746-020-0306-7

Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. *International journal of medical informatics*, *101*, 75–84. https://doi.org/10.1016/j.ijmedinf.2017.02.002

Ilias A., Kanavos, A. & Adamopoulos, M., 2020. "A Medical Decision Support System for the Differential Diagnosis Based on Medical Information Text Mining," *Biomedical Journal of Scientific & Technical Research, Biomedical Research Network+,* vol. 30(5), 23773-23778 (2020). DOI: 10.26717/BJSTR.2020.30.005020

Istepanian, R., Jovanov, E., & Zhang, Y. T. (2004). Introduction to the special section on M-Health: beyond seamless mobility and global wireless health-care connectivity. *IEEE transactions on information technology in biomedicine : a publication of the IEEE*

Engineering in Medicine and Biology Society, 8(4), 405–414. https://doi.org/10.1109/titb.2004.840019

Kadakia, K., Patel, B. & Shah, A. Advancing digital health: FDA innovation during COVID-19. *npj Digit. Med*. 3, 161 (2020). https://doi.org/10.1038/s41746-020-00371-7

Kalla, Dinesh & Samiuddin, Vatsalya. (2020). Chatbot for Medical Treatment using NLTK https://dx.doi.org/10.9790/0661-2201035056

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W. T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebvre, R. C., Mohr, D. C., Murphy, S. A., Quinn, C., Shusterman, V., & Swendeman, D. (2013). Mobile health technology evaluation: the mHealth evidence workshop. *American journal of preventive medicine*, 45(2), 228–236. https://doi.org/10.1016/j.amepre.2013.03.017

The Lancet Rheumatology (2021). Too long to wait: the impact of COVID-19 on elective surgery. *The Lancet. Rheumatology*, 3(2), e83. https://doi.org/10.1016/S2665-9913(21)00001-1

Nguyen, M., Fujioka, J., Wentlandt, K. *et al.* Using the technology acceptance model to explore health provider and administrator perceptions of the usefulness and ease of using technology in palliative care. *BMC Palliat Care* **19,** 138 (2020). https://doi.org/10.1186/s12904-020-00644-8

Peek, N., Sujan, M., & Scott, P. (2020). Digital health and care in pandemic times: impact of COVID-19. *BMJ health & care informatics*, 27(1), e100166. https://doi.org/10.1136/bmjhci-2020-100166

Patrick Ndayizigamiye, Kante, M. & Shingwenyana, S. Jason Warnick (Reviewing editor) (2020) An adoption model of mHealth applications that promote physical activity, *Cogent Psychology*, 7:1, https://doi.org/10.1080/23311908.2020.1764703

Wong, L. E., Hawkins J. E., Langness, S., Murrell, K. L., Iris, P., MD & A. Sammann (2020). Where Are All the Patients? Addressing Covid-19 Fear to Encourage Sick Patients to

Seek Emergency Care. *NEJM Catalyst*, (Figure 1), 1–12. https://doi.org/https://catalyst.nejm.org/doi/abs/10.1056/CAT.20.0193

Websites

Bfarm, (2020), Digitale Gesundheitsanwendungen (DiGA) https://www.bfarm.de/DE/Medizinprodukte/DVG/_node.html

Blümel, M. & Busse, R. (2020), *International Health Care System Profiles Germany,* Commonwealth Fund,. (https://www.commonwealthfund.org/international-health-policy-center/countries/germany).

Chironhealth, (2020), Educate yourself on state telemedicine regulations https://chironhealth.com/telemedicine/regulations/

Europe.eu, (2015), Online consumer reviews The case of misleading or fake reviews https://www.eesc.europa.eu/resources/docs/online-consumer-reviews---the-case-of-misleading-or-fake-reviews.pdf

Jabeen, H., (October 23, 2018), *Stemming and Lemmatization in Python*, Datacamp, https://www.datacamp.com/community/tutorials/stemming-lemmatization-python

Lee, A., (2019), *Why NLP is important and it'll be the future — our future*, towards data science, https://towardsdatascience.com/why-nlp-is-important-and-itll-be-the-future-our-future-59d7b1600dda

Mckinsey & Company, (2020), Report: eHealth Monitor 2020 Deutschlands Weg in die digitale Gesundheitsversorgung – Status quo und Perspektiven https://www.mckinsey.de/~/media/mckinsey/locations/europe%20and%20middle%20east/deutschland/news/presse/2020/2020-11-12%20ehealth%20monitor/ehealth%20monitor%202020.pdf

PWC, (2020), Digital Healthcare https://www.strategyand.pwc.com/de/en/insights/2020/digital-healthcare.html

Roland Berger, (2020), Future of Health – The rise of healthcare platforms https://www.rolandberger.com/en/Insights/Publications/Future-of-Health-The-rise-of-healthcare-platforms.html

Shaaolian, G., (2017), *5 Mobile App Features Your Customers Want,* Forbes https://www.forbes.com/sites/gabrielshaoolian/2017/06/07/5-mobile-app-features-your-customers-want/?sh=fb7ed1e47bf9

Stewart, C., (2020), *Growth in the number of medical apps downloaded during the COVID-19 pandemic by country in 2020**, Statista, https://www.statista.com/statistics/1181413/medical-app-downloads-growth-during-covid-pandemic-by-country/

Tikkanen, R., Osborn R., Mossialos E., Djordjevic A. & Wharton G.A. (2020), *International Health Care System Profiles Germany*, Commonwealth Fund,. https://www.commonwealthfund.org/international-health-policy-center/countries/united-states

United Nations Foundation and Vodafone Foundation. (2009). mHealth for development: The opportunity of mobile technology for health care in the developing world http://www.globalproblems-globalsolutions-files.org/unf_website/assets/publications/technology/mhealth/mHealth_for_Development_full.pdf

WHO, (2018), mHealthNew horizons for health through mobile technologies https://www.who.int/goe/publications/goe_mhealth_web.pdf

Google Play Store Application Websites

Ada – checke deine Gesundheit
https://play.google.com/store/apps/details?id=com.ada.app&hl=de&gl=US

WebMd: Check Symptoms, Rx Savings, & Find Doctors
https://play.google.com/store/apps/details?id=com.webmd.android&hl=de&gl=US

mySugr - Diabetes App & Blutzucker Tagebuch
https://play.google.com/store/apps/details?id=com.mysugr.android.companion&hl=de&gl=US

Doc on Demand
https://play.google.com/store/apps/details?id=com.doctorondemand.android.patient&hl=de&gl=US

K Health | Telehealth, Primary Care & Pediatrics
https://play.google.com/store/apps/details?id=ai.kanghealth&hl=de&gl=US

MyChart
https://play.google.com/store/apps/details?id=epic.mychart.android&hl=de&gl=US

FollowMyHealth®
https://play.google.com/store/apps/details?id=com.jardogs.fmhmobile&hl=de&gl=US

Samsung Health
https://play.google.com/store/apps/details?id=com.sec.android.app.shealth&hl=de&gl=US

Huawei Health
https://play.google.com/store/apps/details?id=com.huawei.health&hl=de&gl=US

Google Fit: Gesundheits- und Aktivitätstracking
https://play.google.com/store/apps/details?id=com.google.android.apps.fitness&hl=de&gl=US

Tabletten Erinnerung für Medikamente und Pillen
https://play.google.com/store/apps/details?id=eu.smartpatient.mytherapy&hl=de&gl=US

Menstruations-Kalender Clue: Perioden & Zyklus-App
https://play.google.com/store/apps/details?id=com.clue.android&hl=de&gl=US

Menstruations-Kalender
https://play.google.com/store/apps/details?id=com.lbrc.PeriodCalendar&hl=de&gl=US

Fabulous – Motivierend!
https://play.google.com/store/apps/details?id=co.thefabulous.app&hl=de&gl=US

Fitbit
https://play.google.com/store/search?q=Fitbit&c=apps&hl=de&gl=US

beurer HealthManager
https://play.google.com/store/apps/details?id=com.beurer.connect.healthmanager&hl=de&gl=US

GooglePay: A safe & helpful way to manage money
https://play.google.com/store/apps/details?id=com.google.android.apps.nbu.paisa.user&hl=de&gl=US

Cash App
https://play.google.com/store/apps/details?id=com.squareup.cash&hl=de&gl=US

PayPal
https://play.google.com/store/apps/details?id=com.paypal.android.p2pmobile&hl=de&gl=US

WorldRemit – Online Geldüberweisung
https://play.google.com/store/apps/details?id=com.worldremit.android&hl=de&gl=US

Acorns - Invest Spare Change

https://play.google.com/store/apps/details?id=com.acorns.android&hl=de&gl=US

Fidelity Investments

https://play.google.com/store/apps/details?id=com.fidelity.android&hl=de&gl=US

Stash: Invest & Build Wealth

https://play.google.com/store/apps/details?id=com.stash.stashinvest&hl=de&gl=US

TrueBill Budget Planner, Bill Tracker and Reminder

https://play.google.com/store/apps/details?id=com.truebill&hl=de&gl=US

Mint: Budget, Bills, Finance & Tax Refund Tracker

https://play.google.com/store/apps/details?id=com.mint&hl=de&gl=US

Goodbudget: Budget & Finance

https://play.google.com/store/apps/details?id=com.dayspringtech.envelopes&hl=de&gl=US

Appendix Codes

The following part presents the most important codes used, in the chronological order of the methodology. Additionally a strawman, showing the hole process will be uploaded of Github under: MaximilianSimmerer

Data Collection

The following packages were installed:

```
import json
import pandas as pd
from tqdm import tqdm
import seaborn as sns
import matplotlib.pyplot as plt
from pygments import highlight
from pygments.lexers import JsonLexer
from pygments.formatters import TerminalFormatter
!pip install google-play-scraper
from google_play_scraper import Sort, reviews, app
%matplotlib inline
%config InlineBackend.figure_format='retina'
sns.set(style='whitegrid', palette='muted', font_scale=1.2)
```

The applications then are implemented into app-packages:

```
app_packages = ['application-URL']
```

Followed by the info scrapping for each individual application:

```
app_infos = []

for ap in tqdm(app_packages):
 info = app(ap, lang='eng', country='US')
 del info['comments']
```

```
 app_infos.append(info)
```

With the info function can be adapted to the language and country one wants to analyze.

To be more precise a helper function for JSON objects was added:

```
def print_json(json_object):
json_str = json.dumps(
json_object,
indent=2,
sort_keys=True,
 default=str
 )
 print(highlight(json_str, JsonLexer(), TerminalFormatter()))
```

Now all valuable information of each application can be identified.
As a next step all available data in the determined language and country was scrapped from the chosen application:

```
app_reviews = []

for ap in tqdm(app_packages):
  for score in list(range(1, 6)):
    for sort_order in [Sort.RATING, Sort.NEWEST]:
      rvs, _ = reviews(
        ap,
        lang='de',
        country='DE',
        sort=sort_order,
        count= 100000 if score == 2 else 100000,
        filter_score_with=score
      )
      for r in rvs:
        r['sortOrder'] = 'Rating' if sort_order == Sort.RATING else 'newest'
```

```
    r['appId'] = ap
  app_reviews.extend(rvs)
```

Now a new data frame with all necessary information was created.

Review rating score & amount analyses

Transforming the column rating in to quarterly average ratings the following function was used:

```
df=df.resample('Q').mean()
```

Transforming the column comments in to quarterly review amount the following function was used:

```
df_F1=df_F1.resample('Q').count()
```

The visual juxtaposition was done with matplotlib and seaborn.

```
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
plt.rcParams['figure.figsize'] = (20,6)
df_F1
df_F2
df_F3
df_F4
df_F5
df_F6
df_F7
df_F8
df_F9
df_F10
```

```
fig = plt.figure()
ax = df_F1.plot()
l = ax.get_lines()
df_F1.plot(ax=ax, color= 'pink')
df_F2.plot(ax=ax, color= 'y')
df_F3.plot(ax=ax, color= 'g')
df_F6.plot(ax=ax, color= 'b')
df_F5.plot(ax=ax, color= 'orange')
df_F4.plot(ax=ax, color= 'grey')
df_F7.plot(ax=ax, color= 'purple')
df_F8.plot(ax=ax, color= 'black')
df_F9.plot(ax=ax, color= 'r')
df_F10.plot(ax=ax, color= 'brown')
```

For each subgroup the visualization was adapted to the required data sets.

Identification Outliers

Clustering reviews based on rating:

```
def sentiment(df_F):
    if df_F['Rating'] == 5:
        return "verypositive"
    elif df_F['Rating'] == 4:
        return "Positive"
    elif df_F['Rating'] == 2:
        return "Negative"
    elif df_F['Rating'] == 1:
        return "verynegative"
    else:
        return "neutrale"

df_F['Sentiment_Specific'] = df_F.apply(sentiment, axis=1)
```

Create new columns including occurrence of rating:

VeryNegative = "|".join(('verynegative', 'kuh', 'beisl'))
Negative = "|".join(('Negative', 'banane', 'apfel'))
Neutrale = "|".join(('neutrale', 'banane', 'wurst'))
Positive = "|".join(('Positive', 'bein', 'nase'))
VeryPositive = "|".join(('verypositive', 'mau', 'pips'))

Put those elements in their own list of objects for easy iteration:

elements = [Negative, Neutrale, Positive, VeryNegative, VeryPositive]

This will create new columns that tell us if an element is in the comment (True or False):

for element in elements:
df_F[element[0:8]] =df_F['Sentiment_Specific'].str.count(element)

Rename the columns for easy use:

df_F.rename(columns = {'verynega' : 'VeryNegative', 'negative': 'Negative', 'neutrale': 'Neutral', 'positive': 'Positive', 'veryposi' : 'VeryPositive'}, inplace = True)
df_F.head()

Counting occurrences in data set:

df_F1=df_F.resample('1m').sum()

Visualization of distribution:

import matplotlib.pyplot as plt

Choose columns in the order to "stack" them:

plt.rcParams['figure.figsize'] = (10,8)

```
df_F1[['VeryNegative', 'VeryPositive', 'Negative', 'Positive', 'Neutral']].plot(kind="bar",
stacked=True)
plt.title("AppName Country Rating Share 2020")
plt.xlabel("Year")
plt.ylabel("Amount")
plt.xlim(0, 13)
plt.ylim(0, 5000)
plt.show()
```

Text analyses of outliers

Intalled packages:

```
import pandas as pd
import numpy as np
from nltk.corpus import wordnet
import string
from nltk import pos_tag
from nltk.corpus import stopwords
from nltk.tokenize import WhitespaceTokenizer
from nltk.stem import WordNetLemmatizer
from nltk.tokenize import word_tokenize
from nltk import PorterStemmer
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import matplotlib.pyplot as plt
```

Text analyses:

```
def get_wordnet_pos(pos_tag):
    if pos_tag.startswith('J'):
        return wordnet.ADJ
    elif pos_tag.startswith('V'):
        return wordnet.VERB
```

```
    elif pos_tag.startswith('N'):
        return wordnet.NOUN
    elif pos_tag.startswith('R'):
        return wordnet.ADV
    else:
        return wordnet.NOUN
```

A function to remove emojis from the reviews

```
def deEmojify(inputString):
    return inputString.encode('ascii', 'ignore').decode('ascii')

def clean_text(text):
    lower text
    text = text.lower()
    Remove Emojis
    text=deEmojify(text)
    Tokenize text and remove puncutation
     text = text.split()
    text = [word.strip(string.punctuation) for word in text.split(" ")]
    Remove words that contain numbers
    text = [word for word in text if not any(c.isdigit() for c in word)]
    Remove stop words
    stop = stopwords.words('english')
    text = [x for x in text if x not in stop]
    Remove empty tokens
    text = [t for t in text if len(t) > 0]
    pos tag text
    pos_tags = pos_tag(text)
    Lemmatize text
    text = [WordNetLemmatizer().lemmatize(t[0], get_wordnet_pos(t[1])) for t in      pos_tags]
    remove words with only one letter
    text = [t for t in text if len(t) > 1]
    join all
    text = " ".join(text)
```

```
    return(text)
```

Clean text data

```
df_F["Comment_clean"] = df_F["Comment"].apply(lambda x: clean_text(x))
```

Word frequency analyses

Divide ratings into positive, negative and neutral segments:

```
def sentiment(df_F):
    if df_F['Rating'] > 3:
        return "positive"
    elif df_F['Rating'] < 3:
        return "negative"
    else:
        return "neutral"

df_F['Sentiment'] = df_F.apply(sentiment, axis=1)
```

Extracting all positive reviews and converting to a list:

```
positive_reviews=df_F.loc[df_F['Sentiment']=='positive','Comment_clean'].tolist()
```

Overview 50 positive comments:

```
positive_reviews[0:50]
```

Frequency counter:

```
from collections import Counter
```

Return words with the highest frequencies:

```
def getMostCommon(reviews_list,topn=40):
    reviews=" ".join(reviews_list)
    tokenised_reviews=reviews.split(" ")
```

```
    freq_counter=Counter(tokenised_reviews)
    return freq_counter.most_common(topn)
```

Visualizing frequency:

```
def plotMostCommonWords(reviews_list,topn=20,title="Common Review
Words",color="blue",axis=None):
```

Default number of words is given as 20

```
    top_words=getMostCommon(reviews_list,topn=topn)
    data=pd.DataFrame()
    data['words']=[val[0] for val in top_words]
    data['freq']=[val[1] for val in top_words]
    if axis!=None:
        sns.barplot(y='words',x='freq',data=data,color=color,ax=axis).set_title(title+" top
"+str(topn))
    else:
        sns.barplot(y='words',x='freq',data=data,color=color).set_title(title+" top "+str(topn))
```

Plot frequency:

```
from matplotlib import rcParams
```

Sets the heigth and width of image:

```
rcParams['figure.figsize'] = 8,6
fig,ax=plt.subplots(1,2)
```

Adjusts the space between the two plots:

```
fig.subplots_adjust(wspace=0.5)
plotMostCommonWords(positive_reviews,40,"PositiveReview Unigrams",axis=ax[0])
plotMostCommonWords(negative_reviews,40,"NegativeReview
Unigrams",color="red",axis=ax[1])
```

Function for bigrams and trigrams:

```
def generateNGram(text,n):
    tokens=text.split(" ")
    ngrams = zip(*[tokens[i:] for i in range(2)]) ## i in range define the length of the word combination
    return ["_".join(ngram) for ngram in ngrams]
```

Visualizing bigrams and trigrams:

```
positive_reviews_bigrams=[" ".join(generateNGram(review,2)) for review in positive_reviews]
negative_reviews_bigrams=[" ".join(generateNGram(review,2)) for review in negative_reviews]

rcParams['figure.figsize'] = 8,6
fig,ax=plt.subplots(1,2)
fig.subplots_adjust(wspace=1)
plotMostCommonWords(positive_reviews_bigrams,20,"Positive Review Bigrams",axis=ax[0])

plotMostCommonWords(negative_reviews_bigrams,20,"Negative Review Bigrams",color="red",axis=ax[1])
```

Merging Data sets into one

Transform all data sets into a list:

```
df_list = [df_F1, df_F2, df_F3 , df_F4, df_F5,df_F6, df_F7, df_F8 , df_F9, df_F10 ]
df_list
```

Transform list into merged dataframe:

```
df_Healthcare_US = pd.concat(df_list)
df_Healthcare_US
```

Save new merged data frame as .csv file:

df_Healthcare_US.to_csv('review_Healthcare_US_Combined.csv', index=None, header=True)

Key feature analyses

Dividing data set by sentiment of rating score:

Bad Reviews:
df_F['Rating'] = df_F.Rating.astype(str)
df_F.drop(df_F.loc[df_F['Rating'] > '2.5'].index, inplace=True)

Good Reviews:
df_F['Rating'] = df_F.Rating.astype(str)
df_F.drop(df_F.loc[df_F['Rating'] < '3.5'].index, inplace=True)

Dictionary for Finance key feature analyses:

PerformanceExpectancy = "|".join(('waste',
'time','performance','helpful','help','useful','pleasant','friendly','offer','accomplished','issue','correctly',

'correct','promising','professional','disappointed','useless','update','accessable','accessible','acute',

'adequate','approachable','appropriate','certified','cheap','cheaper','cheapest','cheaply','confuse',

'detailed','detect','detective','determine','develop','developer','development','deviate','device',

'difficult','difficulty','disappear','disappeared','disappoint','disappointed','disappointing','disconnect',

'discount','discover','error','excellence','excellent','exceptional'

,'exceptionally','expectation','fair','fairly','fake','function','functional','functionality',

'functioning','helpful','helpfulness','helplessness','helps','income','inexpensive','innovation','innovative',

'insightful','irrelevant','irresponsible','knowledge','knowledgeable','misdiagnosed','monitor','personal',

'personalized','personally','privacy','recommend','recommendable','recommendation','recommended','regular',
'regularly','regulate','regulation','resource','resourceful',
'satisfactory','satisfied','satisfy', 'secure',
'uncontrolled','usefulness','useless','uselessly'))

EffortExpectancy = "|".join(('flexible', 'reconnect', 'simple', 'clear', 'understand', 'easy',
'complicated', 'accurate', 'experience', 'precise', 'pleasant', 'accomplished', 'wait',
'effective','efficiency','efficient','efficiently','effort', 'execute', 'familiar' , 'fast','faster',
'financial', 'friendly', 'frustrate','frustrated','frustrating' ,'frustration',
'hard','harder','hardly','hardship','hassle','haste', 'immediate'
,'immediately','immensely','immitately', 'inaccurately','inadequate', 'incompetent'
, 'inconsistent','inconvenience','incorrect','incorrectly',
'info','inform','informality','information','informational','informative','informed',
'instruct','instruction','integration', 'interface', 'interact','interaction','interactive',
'lasted','late','latently','later', 'mishandle','mislead','misleading','misleads',
'misread','problem','problems', 'process'
,'professional','professionalism','professionally','relatable','relate','related','reply'
,'request','respond','responded','responds','response','responsive','responsiveness'
,'retrying','return','setting','simple','simplest','simplicity','simplify','simply'
,'software','speed','standard','straightforward','subscription','unclear'
,'understand','understanding','understands','understood','unfromative','uneasy'
,'unseen','unusable','unusefully'))

SocialInfluence = "|".join(('advise', 'tipp', 'recommend',
'recommendation','men','women','husband', 'wife', 'brother', 'sister' , 'son', 'daugther', 'family',
'mum', 'dad', 'mother', 'father', 'grandfather' , 'grandmother', 'granddaughter', 'grandson',
'friends', 'friend' 'bestfriend', 'collgeaus', 'boss','neighbor', 'nurse', 'overtime','overuse',
'social','society', 'internet', 'boyfriend', 'girlfriend', 'partner', 'facebook', 'twitter', 'instagram'))

ProcessExpectancy = "|".join(('transaction', 'payment','charge', 'bank', 'finance', 'platform',
'transfer', 'slow', 'instal','deinstall','phone',
'crash','update','upgrade','uninstall','uninstalled','uninstalling','unreachable','unreliable',
'install','installed','security','unauthorized','fail','fails','failure', 'service', 'customer service' ,
'customer support','help', 'mobile service', 'mobile', 'data', 'version', 'account', 'sign', 'signin',
'login','verification','verify','works','subscription','unclear','synchronization', 'error',
'customer', 'feature','function','guarantee', 'guidance','guide','guideline', 'hotline', 'launch',
'legal' , 'weight', 'management', 'step', 'count', 'heart', 'rate', 'track', 'tracker', 'last update',
'watch', 'face', 'late version', 'update device', 'magic watch', 'sync', 'password', 'change',
'device','delete','maintenance', 'platform','reconnected','reconnecting', 'reinstall','reinstalling',
'update', 'version'))

Put those elements in their own list of objects for easy iteration
elements = [PerformanceExpectancy, EffortExpectancy, SocialInfluence, ProcessExpectancy]

This will create new columns that tell us if an element is in the comment (True or False)
for element in elements:
    df_F1[element[0:3]] =df_F1['Comment_clean'].str.count(element)

Rename the columns for easy use
df_F1.rename(columns = {'was': 'PerformanceExpectancy', 'fle': 'EffortExpectancy', 'adv':
'SocialInfluence', 'tra': 'ProcessExpectancy'}, inplace = True)
df_F1.head(50)

Dictionary for German key feature analyses:

PerformanceExpectancy = "|".join(('zufrieden','brauchbar', 'letzte version', 'version', 'neue
version', 'neues update', 'enttäuschend', 'zufrieden', 'zufriedenstellend', 'unbrauchbar',

'nützlich', 'hilfreich', 'zufrieden', 'nützlich','sicherheit', 'privatssphäre', 'probleme', 'freundlich', 'unfreundlich', 'unzuverlässig' , 'professinell', 'unprofessionell', 'beeindurckend', 'unbeeindruckend', 'motivierend' , 'schnell', 'praktisch', 'genau', 'präzise', 'zuverlässig'))

EffortExpectancy = "|".join(('übersichtlich','einstellungen','änderung','änderungen', 'änderungsmöglichkeiten' ,'einfach zu bedienen', 'app lässt sich einfach bedienen', 'einloggen' , 'account erstellen', 'übersichtlich', 'verständlich', 'installation', 'unverständlich', 'unnötig', 'simpel', 'umständlich', 'nervig', 'kundenservice' , 'service', 'kontakt', 'erreichbar', 'erreichbarkeit', 'angenehm', 'connect' , 'verbindung', 'verbinden', 'unkompliziert', 'kompliziert', 'zeitaufwending', 'schnell'))

SocialInfluence = "|".join(('gut gestaltent','freunde', 'freund', 'freundin', 'mutter', 'vater', 'eltern' , 'großmutter', 'großvater', 'oma', 'opa', 'schwester', 'tochter', 'bruder', 'sohn', 'enkel', 'enkelin', 'verwandte', 'bekannte', 'bekannter', 'verwandtschaft', 'tante , 'onkel', 'kollegen', 'kollegin', 'kollege' , 'familie', 'empfhelung', 'empfohlen', 'weiter empfohlen', 'tipp', 'rat', 'ratschlag'))

ProcessExpectancy="|".join(('installieren','installiert','sicherheit','unautorisiert','fehlschlagen','scheitern' ,'versagen','service','kundendienst' ,'kundendienst','hilfe','mobilerservice','mobil','Daten','version','konto','unterschreiben','anmelden','einloggen', 'verifizierung','verifizieren', 'verifikation', 'funktioniert','abonnement','unklar' ,'synchronisation','fehler','kunde',

'feature','funktion','garantie','anleitung','leitfaden','richtlinie','hotline','launch','legal'

,'gewicht','management','schritt','zählen','herz','rate','verfolgen','tracker','letzte aktualisierung', 'watch', 'face', 'late version', 'update device', 'magic watch', 'sync', 'gewichtsverwaltung' ,'update', 'letztes update', 'seit dem letzen update', 'passwort', 'anmelden', 'neuste,' 'wartung', 'plattform','wiederverbunden','wiederverbinden', 'neu installieren','neuinstallieren'))

Put those elements in their own list of objects for easy iteration

elements = [PerformanceExpectancy, EffortExpectancy, SocialInfluence, ProcessExpectancy]

This will create new columns that tell us if an element is in the comment (True or False)

```
for element in elements:

    df_F2[element[0:3]] =df_F2['Comment_clean'].str.contains(element)
```

Rename the columns for easy use

```
df_F2.rename(columns = {'zuf': 'PerformanceExpectancy', 'übe': 'EffortExpectancy', 'ins':
'ProcessExpectancy', 'gut': 'SocialInfluence'}, inplace = True)
df_F2.head()
```

Dictionary for U.S. key feature analyses:

```
PerformanceExpectancy = "|".join(('waste',
'time','performance','helpful','help','useful','pleasant','friendly','offer','accomplished','issue','co
rrectly',
'correct','promising','professional','disappointed','useless','update','accessable','accessible','ac
ute',
'adequate','approachable','appropriate','certified','cheap','cheaper','cheapest','cheaply','confu
se',
'detailed','detect','detective','determine','develop','developer','development','deviate','device',
'difficult','difficulty','disappear','disappeared','disappoint','disappointed','disappointing','disco
nnect',  'discount','discover','error','excellence','excellent','exceptional'
,'exceptionally','expectation','fair','fairly','fake','function','functional','functionality',
'functioning','helpful','helpfulness','helplessness','helps','income','inexpensive','innovation','inn
ovative',
'insightful','irrelevant','irresponsible','knowledge','knowledgeable','misdiagnosed','monitor','p
ersonal',
'personalized','personally','privacy','recommend','recommendable','recommendation','recomm
ended','regular',  'regularly','regulate','regulation','resource','resourceful',
'satisfactory','satisfied','satisfy',  'secure',  'uncontrolled','usefulness','useless','uselessly' ))
```

EffortExpectancy = "|".join(('flexible', 'reconnect', 'simple', 'clear', 'understand', 'easy',
'complicated', 'accurate', 'experience', 'precise', 'pleasant', 'accomplished', 'wait',
'effective','efficiency','efficient','efficiently','effort', 'execute', 'familiar', 'fast','faster', 'financial',
'friendly', 'frustrate','frustrated','frustrating' ,'frustration',
'hard','harder','hardly','hardship','hassle','haste', 'immediate'
,'immediately','immensely','immitately', 'inaccurately','inadequate', 'incompetent'
, 'inconsistent','inconvenience','incorrect','incorrectly', 'info','inform'
,'informality','information','informational','informative','informed',
'instruct','instruction','integration', 'interface', 'interact','interaction','interactive',
'lasted','late','latently','later' , 'mishandle','mislead','misleading','misleads',
'misread','problem','problems', 'process',
'professional','professionalism','professionally','relatable','relate','related','reply'
,'request','respond','responded','responds','response','responsive','responsiveness'
,'retrying','return','setting','simple','simplest','simplicity','simplify','simply'
,'software','speed','standard','straightforward','subscription','unclear'
,'understand','understanding','understands','understood','unfromative','uneasy'
,'unseen','unusable','unusefully'))

SocialInfluence = "|".join(('advise', 'tipp', 'recommend',
'recommendation','men','women','husband', 'wife', 'brother', 'sister', 'son', 'daugther', 'family',
'mum', 'dad', 'mother', 'father', 'grandfather', 'grandmother', 'granddaughter', 'grandson',
'friends', 'friend' 'bestfriend', 'collgeaus', 'boss','neighbor', 'nurse', 'overtime','overuse',
'social','society', 'internet', 'boyfriend', 'girlfriend', 'partner', 'facebook', 'twitter', 'instagram'))

ProcessExpectancy = "|".join(('instal','deinstall','phone',
'crash','update','upgrade','uninstall','uninstalled','uninstalling','unreachable','unreliable',
'install','installed','security','unauthorized','fail','fails','failure', 'service', 'customer service'
, 'customer support','help', 'mobile service', 'mobile', 'data', 'version', 'account', 'sign', 'signin',
'login', 'verification','verify','works','subscription','unclear','synchronization', 'error',
'customer', 'feature','function','guarantee', 'guidance','guide','guideline', 'hotline', 'launch',
'legal' , 'weight', 'management', 'step', 'count', 'heart', 'rate', 'track', 'tracker', 'last update',
'watch', 'face', 'late version', 'update device', 'magic watch', 'sync', 'password', 'change',

'device','delete', 'maintenance', 'platform','reconnected','reconnecting', 'reinstall','reinstalling', 'update', 'version'))

Put those elements in their own list of objects for easy iteration

elements = [PerformanceExpectancy, EffortExpectancy, SocialInfluence, ProcessExpectancy]

This will create new columns that tell us if an element is in the comment (True or False)

```
for element in elements:
    df_F1[element[0:3]] =df_F1['Comment_clean'].str.count(element)
```

rename the columns for easy use

```
df_F1.rename(columns = {'was': 'PerformanceExpectancy', 'fle': 'EffortExpectancy', 'adv': 'SocialInfluence', 'ins': 'ProcessExpectancy'}, inplace = True)
df_F1.head(50)
```

Tables Key Feature Distribution

Medicine Monitor U.S.

	Negative	Positive	Neutral	VeryNegative	VeryPositive
PerformanceExpectancy	520	6672	1243	1316	29694
EffortExpectancy	321	3866	678	699	16387
SocialInfluence	128	1128	297	288	4744
ProcessExpectancy	778	5733	1313	1992	25576

Telehealth U.S.

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	760	610	239	843
Positive	3742	5627	1486	3146
Neutral	858	773	298	813
VeryNegative	4052	2533	1130	4317
VeryPositive	26798	40268	10572	18426

Health & Fitness U.S.

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	28099	11180	3332	59279
Positive	32037	7965	2879	35288

Neutral	31293	11413	3871	58110
VeryNegativ	60434	21084	7048	125674
VeryPositiv	49038	7557	2048	23179

Medicine Monitor Germany

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	533	126	48	386
Positive	6563	845	370	1143
Neutral	1152	192	95	525
VeryNegative	557	109	71	694
VeryPositive	47200	5013	1221	4395

Health & Fitness Germany not included in U.S.

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	3241	1305	291	6956
Positive	5244	1025	275	3542
Neutral	3359	1065	279	4977
VeryNegative	8511	3695	918	21585
VeryPositive	9037	1119	380	3085

Health & Fitness Germany included in U.S.

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	3021	882	286	6697
Positive	5332	736	407	4206
Neutrale	3395	832	359	5718
VeryNegative	6559	2237	568	16729
VeryPositive	8334	697	269	2947

Transaction Finance

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	35655	14437	8988	52177
Positive	51120	19712	6206	26875
Neutral	35018	11085	6409	35443
VeryNegative	73719	38788	22725	135423
VeryPositive	224363	131423	22888	66414

Trading Finance

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	4732	2782	1246	8008
Positive	14159	7352	2453	9265
Neutral	5987	3264	1453	8050
VeryNegative	12396	8319	3803	27572
VeryPositive	68145	34595	10842	24184

Budget Planner

	PerformanceExpectancy	EffortExpectancy	SocialInfluence	ProcessExpectancy
Negative	6240	2576	646	10195
Positive	17521	5630	1999	17928
Neutral	7076	2609	796	10688
VeryNegative	10997	4922	1229	19245
VeryPositive	43409	15764	5970	35207