

How Patients rate Physicians? Analyzing key factors of patient satisfaction using German Physician-Rating Platform

Social Media and Business Analytics (Research Project)

Research Paper

Divya Mugthihalli Devaprasad
divya.mugthihalli.devaprasad@uni-potsdam.de
Matrikel-Nr. 805132

Sandeep Uprety
uprety@uni-potsdam.de
Matrikel-Nr. 804982

ABSTRACT

The health care sector has evolved exponentially with the growth of the internet. Customer-oriented health care management and patient satisfaction have become an important key factor for physicians to attract patients. In the current competitive environment, platforms such as Online health care (OHC), medical crowd-sourcing, Physician-Rating Websites (PRWs) are gaining popularity with time. Unlike the past where physician popularity was witnessed via word of mouth, at present several online search engines enable patients to locate physicians around them, based on several criteria such as patient's illness type and physician's expertise. Patients satisfaction influences the selection of a physician and leads to higher patient retention and better reviews. In general, positive reviews tend to increase the patient's tendency to follow the advice given by physicians. In recent years PRW has emerged in the health care sector and is increasingly used by patients. Patients' involvement is either through posting an evaluation to give feedback to other patients based on their experience with a physician or selecting other physicians based on posted reviews. The growth of PRW has opened many windows to analyze the patient's satisfaction and its key drivers. PRW data enables satisfaction analysis on the individual level of physicians and as well as implications on the overall level.

Keywords

Online physician rating, Online health care, patient satisfaction, Physician rating platform.

1. INTRODUCTION

In recent years many health care systems have implemented quality of care improvement strategies (Lebrun, et al., 2013), regardless of which lack of quality persists (McGlynn, et al., 2003). Patients are generally not aware of the existing quality differences, the reason being less exposure to publicly reported information on the quality of health care providers. Establishing a transparent mode of communication is a single source of remedy and achieving this has been a major challenge (Porter & Guth, 2012). Transparency in information will tend to increase the overall quality and drive patients to better and reliable health care providers (Faber, et al., 2009) and on the other side it will also motivate the physicians to provide quality service (Fung, et al., 2008). As a result, public reporting (PR) instruments have been set up in many countries (Mukamel & Mushlin, 1998). These instruments help in assessing the quality of care offered by health care providers (Lagu, et al., 2010).

One such trend in the PR movement is the use of PRWs (Emmert, et al., 2012). PRW is an online platform which enables the patient to rate their respective physician. Patients can post their comments or provide feedback to other patients based on their experience with the physician and at the same time read the feedbacks provided by other patients before approaching a physician (Hanauer, et al., 2014). Patients are very particular when it comes to the decision-making process concerning their medical care in the context of changing patient-physician relationships (Zeithaml, et al., 1993). Internet is one of the major source providers of online health care information (Fox, 2014), PRWs are inclined towards providing information related to the quality of care from a patient's point of view (Verhoef, et al., 2014).

From a physician's perspective, PRWs are important because it is a platform that provides information on the physician's service quality and this can be made publicly available. This factor will boost patient satisfaction and help in generating positive word of mouth (Emmert & Meier, 2013). On the other hand, patient evaluation on the physician over PRWs provides a scope of improvement for physicians' service quality. The rise of PRWs offers new platforms to analyze patient satisfaction and its key drivers (Gill & White, 2009).

There exists a large amount of evidence related to traditional PR instruments but only a few types of research have addressed PRWs. A systematic review conducted has identified 9 articles related to PRWs which are published in peer-reviewed journals (Emmert, et al., 2013). Emmert & Meier (2013) in their paper, have mentioned that these articles investigate the number, distribution, and trend of the evaluations on PRWs. Most of the investigations assessed ratings for a (non) random sample of physicians. Additionally, there is no published evidence available that analyzes the characteristics of the patients who provide ratings (Emmert & Meier, 2013).

Existing studies on key factors of patient satisfaction are usually based on small sample sizes (Davis, et al., 2017). From a perspective of the key factor, analyzing data from PRWs facilitates analysis of data on a large scale, i.e., information related to patients and physicians, and also helps to identify how certain specific service attributes contribute to overall patient satisfaction and patients' behavioral intentions.

Thus, the main purpose of our study was to conduct a key driver analysis through hypothesis testing and using a regression model applied to large-scale data from a PRW. In this context, our paper outlines the analysis we conducted on physician evaluations posted on the German PRW-Jameda, from 2008 to 2020. In our research, we provide a descriptive analysis of (1) both physician and patient characteristics, and (2) the number, distribution, and results of the ratings. Analytical analyses were applied to assess (3) the impact of physician and patient characteristics on the overall performance measure, and (4) to get insight on patient review about a physician.

The ability to identify how and to which extent different service attributes contribute to patients' overall satisfaction with a physician is of high importance for the physician (Gill & White, 2009). Hence, the findings from our study will strengthen the understanding of patient satisfaction and contribute to the body of knowledge in the health care sector. Our paper addresses the following research question:

What are the key factors that influence the patient's satisfaction towards a physician?

The remaining paper is structured as follows. The introduction is followed up by Theoretical background and later by Methodology. The methodology section explains the data portion highlighting the data collection and pre-processing part. In the later sections, we present the results of our findings followed by the conclusion and scope for future research.

2. THEORETICAL BACKGROUND

This section outlines the theories that support and form a foundation for our research study. For better understanding we have divided our theoretical background into the following sections:

2.1 Usage and Usefulness of Physician-Rating Websites

Several research papers have been published worldwide by various researchers to investigate PRWs in countries such as Germany (Emmert, et al., 2013), Switzerland (Rothenfluh & Schulz, 2018), the Netherlands (Damman, et al., 2010), and China (Li, et al., 2018). Rothenfluh & Schulz (2018) in their paper have identified and analyzed 143 PRWs in both English and German-speaking countries. PRWs which provide factual information such as opening hours and geographical location of a physician (Fischer & Emmert, 2015), also provides a platform for patients to evaluate a particular physician and review the physician based on their experiences such as quality of care they receive during their time of appointment. PRWs enable patients to post their content in the form of comments and reviews (Greaves, et al., 2012). According to patient's perspectives, PRWs are a convenient tool to share their experiences which they experience during the time of medical care (Lagu, et al., 2010). Compared with other rating websites related to other domains of day to day-to-day life (McLennan, et al., 2018), usage of PRWs is on the rise (Terlutter, et al., 2014).

There are few points of criticism that are attached towards PRWs. The work is inconsistent when it narrows down to the link between the information available on the rating sites and the quality of care provided (Verhoef, et al., 2014). Patients though feel challenged to judge a physician's competence (Emmert, et al., 2014), but do review the physician's competence using the PRWs platform (Rothenfluh & Schulz, 2017).

Several Text mining approaches (Li, et al., 2018) have been used to identify and analyze what patients express in free-text comments. These methods seldom help in identifying only the most frequently mentioned aspects in the comments while on the other hand, they neglect the minority comments which the reviewers have mentioned (Hao & Zhang, 2016). In addition, Hao & Zhang (2016) also express in their paper that people may often avoid posting negative comments on the platform due to privacy concerns. Thus, some PRWs don't rely on free-text responses but rather on rating scales as an alternative. We in our paper have selected PRW (Jameda) where both free-text comments and rating scale are used.

Holliday, et al., (2017) in their paper have addressed that there are mainly two different types of PRWs based on data collection. One being *Independent websites*, such as Healthgrades (Healthgrades, 2019), which are run by private companies and nurtured by crowdsourcing data and other being *health system websites*, i.e., PRWs that can be established by health systems, which gather information from patients in terms of ratings when they visit hospital or physician (Holliday, et al., 2017)). The PRWs we use for our research can be categorized into Independent websites, as it a German search engine in the health care sector that also provides a platform for physician review (Jameda, 2007)

The information available on PRWs are data in numerical rating scales which can be valuable for different stakeholders. Information provided on PRWs should be of interest not only for patients but also for physicians and other health care providers.

2.2 A key driver analysis of patient satisfaction using PRW

A survey conducted on a clinic in Minnesota, USA revealed that 28.1% of patients agreed that a positive review on PRW would influence them to seek consultation from that physician and 27% indicated that a negative review would cause them to not seek consultation from the physician (Burkle & Keegan, 2015). On the other hand, another survey in the social network highlights that a physician's expertise and reputation were the key driving factors for consultation with that particular physician. (Schaarschmidt & Ivens, 2017). With the emergence of PRW, there arises a new platform that helps in analyzing the satisfaction of patients and the key driving force. Patient satisfaction has become an important factor for physicians, towards the goal to attract patients in an increasingly competitive environment (Otani, et al., 2009). PRWs provide a platform for the patients to evaluate their physicians and review them based on the quality of care they receive during an appointment. Concerning the patients' perspective, PRWs are a convenient method to share information about the medical service they have received (Lagu, et al., 2010).

As patients' expectations often prove to be idle over time, individual satisfaction is measured based on perceived performance. In our study, we shall follow this argument by assuming that patients form their overall satisfaction towards the physician based on the service provided and quality of service and not by comparing with the patient expectations (Leisen & Hyman, 2004). Furthermore, as per the study conducted by Wilkie & Pessemier (Wilkie & Pessemier, 1973), we shall use a multi-attribute model (Ittersum, et al., 2007) where we take into account the linear and nonlinear effects of perceived service quality on overall evaluation i.e., service attributes can have a diminishing, constant, or increasing returns. Our study also stands unique by the chosen dataset and the evaluation attributes available on the considered platform.

In this paper, we took an inductive approach to identify the relations between physician attributes and Patient satisfaction by evaluating the key factors influencing patient satisfaction towards the physician. The core contribution of our research was to use and analyse the data available on the Jameda website and establish a relationship between these variables. Our research focused on performing hypothesis testing using the Statistical analytical method - "Median tests" followed by carrying out regression analysis. Finally, based on the coefficients from the tests, we accept/reject the hypothesis at a certain significance level. The following conceptual framework gives an overview of our research study.

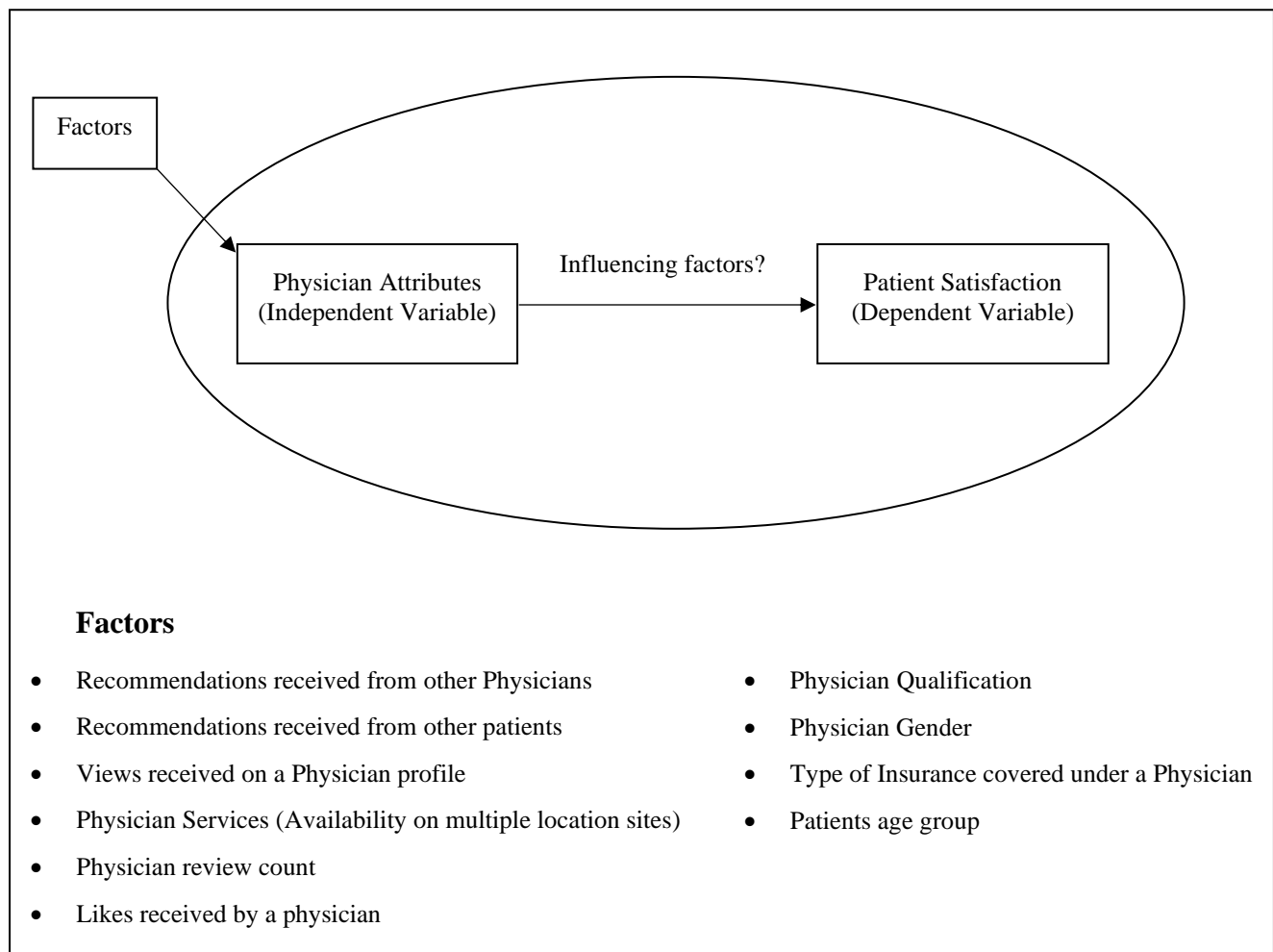


Figure 1: Conceptual Framework

3. METHODOLOGY

This section describes the process of data collection and the methodology followed to answer our research question.

3.1 Approach

The approach followed for our research is shown in a general framework schematized in Figure 2. Physician data and patient reviews were collected from the Jameda website database. Later, from the available data, a specific sample space was selected for our research. Later, Statistical analysis is carried out to establish a relationship between independent variables (Physician attributes) and dependent variable (patient satisfaction) and also to define the hypothesis. The research is then followed by hypothesis testing and regression analysis to deduce the coefficient value. In addition to this, we form a *Word Cloud*, to understand what patients frequently post about their physicians. Finally, based on the results of the coefficient value we accept or reject the hypothesis, which leads us to concrete evidence on answering our research question.

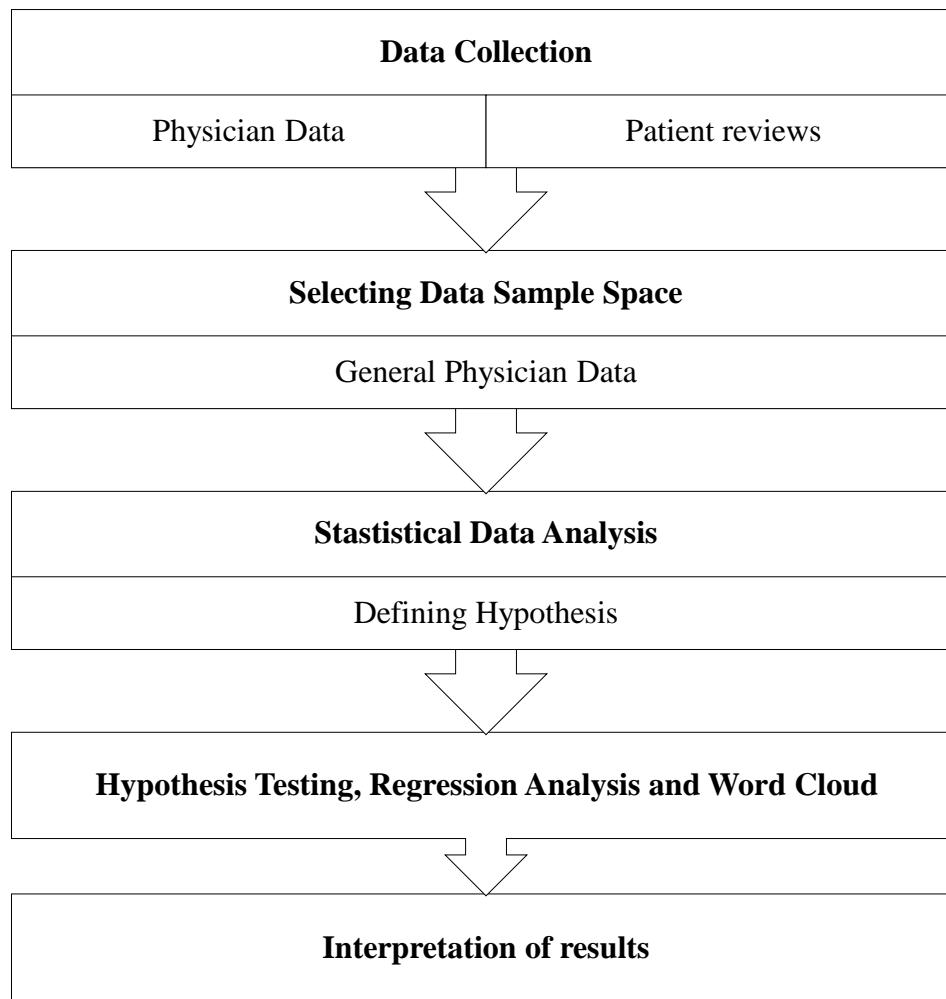


Figure 2: Approach towards research

3.2 Data Collection

Our study is based on the analysis of 815935 physicians who are registered in German PRW, Jameda from 2008 to 2020. Out of 815925 registered physicians, only 345426 physicians have been graded which accounts for 42% of the total data set. For our research, we have considered this 42 % of the data population, which includes a total of 5172248 reviews on these graded and registered physicians. It is found that the highest number of registered physicians in Jameda belongs to a group of “General physicians”, hence this data set of physicians form a basis for our study. Out of 345426 graded and registered physicians, 43405 belong to the category of General physicians (Male and Female) and the patient review count for this physician sample space is 468558. The following graphs indicate the distribution overview of physicians based on physician practice type along with their respective review count. The term “Others” here is referred to the physicians who fall under “*different Practitioner types*” other than General Physicians. The distribution overview of “others” is shown in Figure 6.

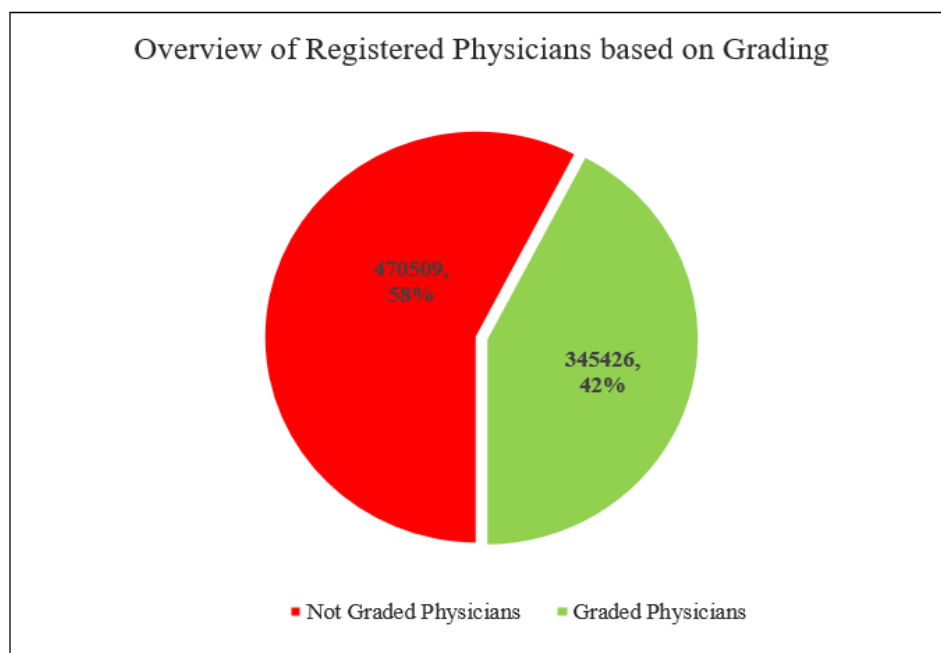


Figure 3: Overview of Registered Physician based on Grading

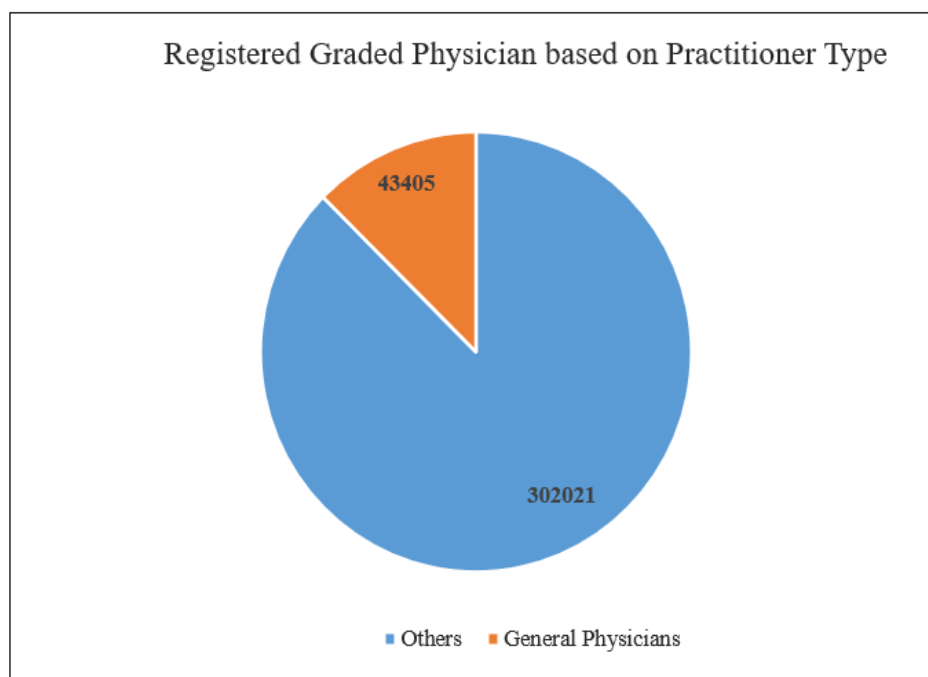


Figure 4: Registered Physician based on Practitioner Type

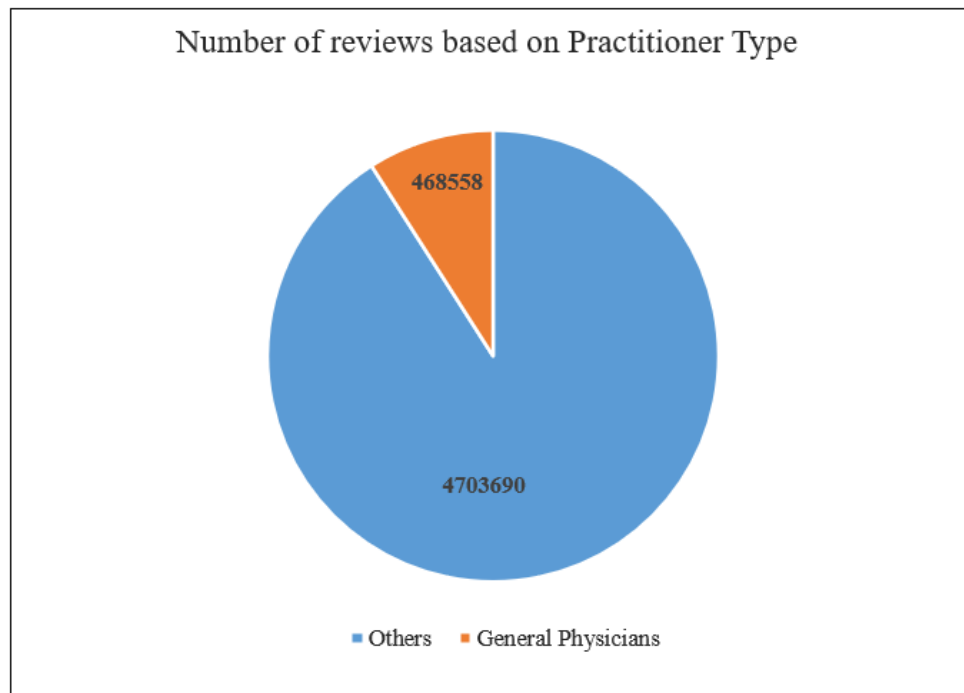


Figure 5: Number of reviews based on Practitioner Type

The dataset contained information related to a physician such as physicians' gender and area of specialty along with the patient's information such as patients' age and type of insurance the patient is insured under. In addition to that, the data set included the results of the physician rating for all compulsory and optional questions. The physician rating system on Jameda consists of five compulsory questions which are rated according to the grading system followed in German schools, i.e., ratings are given on the scale of 1-6 (1=very good; 2=good; 3=satisfactory; 4=fair; 5=deficient; and 6=insufficient) (Geraedts, et al., 2012)

The five compulsory questions are as follows:

1. Is the patient satisfied with the treatment offered by the physician?
2. Does the physician have good knowledge of the type of illness and the treatment?
3. Has the physician built a good trust with his/her patients?
4. Has the physician given sufficient time to the patient's concerns?
5. Is the physician-friendly with his/her patients?

A mean score ("overall grade") is calculated, based on the results of the above five questions. The data set also consisted of other 12 optional questions for answering but these are not addressed in this paper. Our research focused on Jameda because it is considered to play a significant role in the German PRW movement for the following reasons:

- (1) From a patient's perspective, Jameda is the PRW to which a patient is most likely to be referred (Emmert, et al., 2012)
- (2) Jameda is ranked highest in traffic among German PRWs (Jameda, 2007)
- (3) So far, among German PRWs, Jameda consists the largest number of ratings (Emmert, et al., 2012)

3.3 Selecting Data Sample Space

A sample space represents the entire population. Defining a sample space facilitates working on the selected sample data conveniently thereby enabling to analyze and derive accurate results more quickly and efficiently (Bhandari, 2020). As explained in the previous section, from the given database it is found that the number of highest registered physicians on the Jameda website come under the category of General Physicians and hence we believe that selecting this sample space will help us address our research question. The below graph gives an overview of the top ten categories of registered physicians and General physicians being the highest.

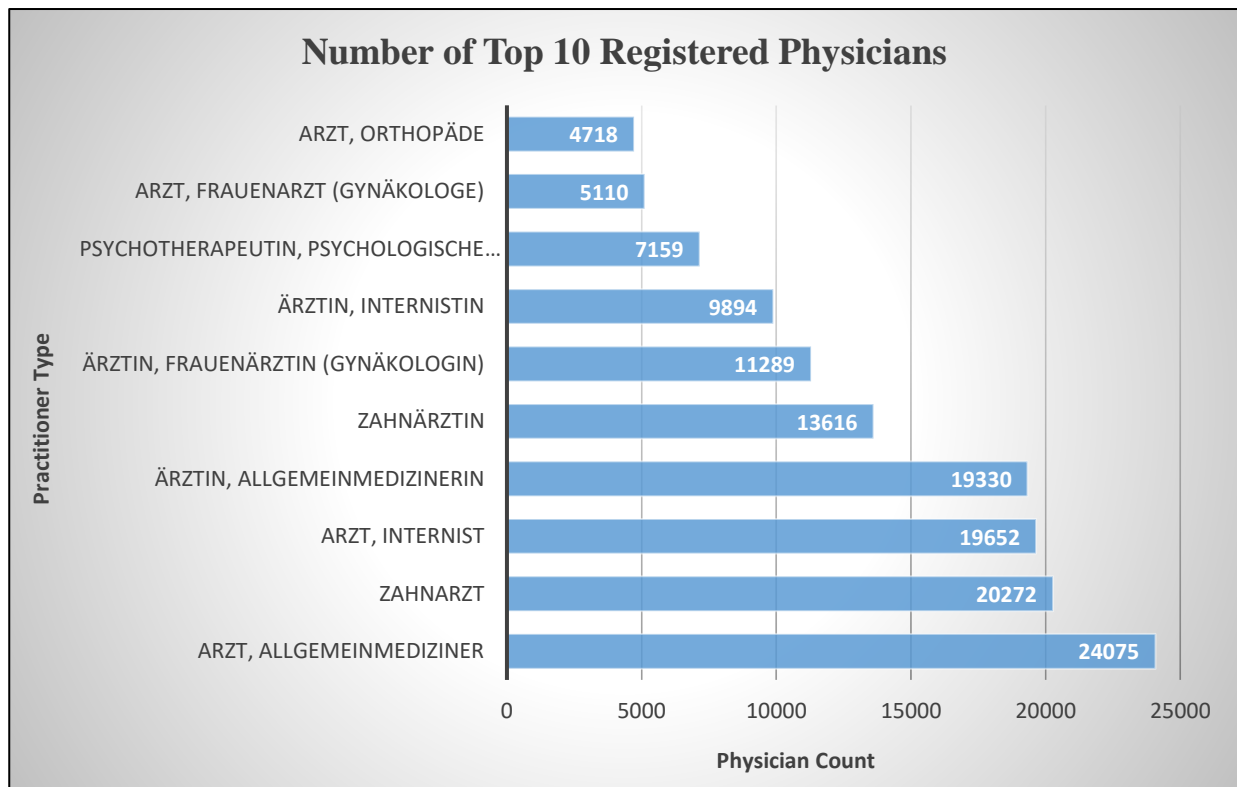


Figure 6: Number of Top 10 Registered Physicians

3.4 Statistical Data Analysis

Once we have defined our sample space, to quantify the data, we perform Statistical Data analysis on our selected sample space. We begin our analysis by comparing the overall rating received by a physician based on physician attributes and patient characteristics. The following tables give an overview of factors which are considered for physician attributes and patient characteristics.

Sl. No	Physician Attributes
1	Physician Popularity
2	Physician Gender
3	Physician Qualification
4	Physician Services

Table 1: Physician Attributes

Sl. No	Patient Characteristics
1	Type of Insurance covered under a Physician
2	Patient age

Table 2: Patient Characteristics

The following factors define physician popularity:

- Recommendations received from other Physicians
- Recommendations received from other patients
- Physician review count
- Likes received by a physician
- Number of views received on a physician profile

Table 3 gives an overview of the compared variables of physician and patient attributes. A hypothesis test is carried out individually between the attributes (patient and physician) and their respective variables to identify if there exists a significant difference concerning overall ratings received by a physician or overall ratings given by a patient.

Sl. No	Attributes (Physician and Patient)	Compared Variables
1	Physician Gender	Male Female
2	Physician Qualification	Highly Qualified Qualified
3	Physician Services	Single Location consultation Multiple Location consultation
4	Type of Insurance covered under a Physician	Private Insurance Statutory Insurance Type not available
5	Patient age group	Younger patients (age < 30) Older patients (age > 50)

Table 3: Compared Variables

3.5 Hypothesis Testing, Regression Analysis, and Word Cloud

In our research, we have conducted the statistical analysis in python using the python *scipy stats* package. Since our data is non-parametric and the grades are ordinal, the median test is used for hypothesis testing where the comparison is made between the median of the variables. The assumption of the statistical test is called Null Hypothesis (H_0) and Alternative Hypothesis (H_1). We make the following assumptions for our Median test.

H_0 : The medians of the compared variables are the same

H_1 : The medians of the compared variables are different

The median test thus carried out on the compared variables, results in p-values. This quantity is further used to interpret the tested Hypothesis (Hernandez, 2017).

Regression analysis is a set of statistical processes and is mainly used for two conceptually distinct purposes. First, it is used for prediction and forecasting. Second, for estimating the relationship between the dependent and independent variable(s) in a fixed dataset (Kang, 2013). In our research, we aim to identify the significant relationships between physician attributes and patient satisfaction. Additionally, the regression model also indicates the strength of the impact of multiple independent variables (physician attributes) on the dependent variable (patient satisfaction).

For our study, we conduct regression analysis using the linear model from the python sklearn package. As the first step for Regression analysis, we consider only the attributes which are required for our analysis. Next, we look for missing values and categorical attributes. Missing data occur in almost all research. Missing data can reduce the statistical power of a study and can produce biased results, leading to invalid conclusions (Kang, 2013). The most common method to approach the missing data is simply omitting those cases which contain the missing data and analyze the remaining data. This approach is known as the complete case (or available case) analysis or listwise deletion (Kang, 2013). The table below illustrates the missing valued attributes and the technique used to handle them.

Attributes with missing values		Number of missing values	Percentage	The technique used to handle missing values
Patient Data	age of patient	73334	46.92	considered only the older and younger patient population
	insurance type of patient	81092	51.88	omit the cases with the missing data
Physician Data	designation of physician	38819	24.83	omit the cases with the missing data
	recommendations from patients	5136	3.29	Imputed with value 0
	likes on physician profile	4794	3.07	Imputed with value 0

Table 4: Missing value attributes and techniques used

Furthermore, we look for the categorical variables. These are the variables that take a finite number of categories or groups. It is required to encode it to numerical values before we fit the linear regression model (Volpi, 2020). The table below illustrates the categorical attributes and the technique used to encode them.

Categorical attributes	Type	Values	Encoding technique used	Values after encoding
Gender of physician	Nominal	“Male”, “Female”	one hot encoding	0,1
Age of the patient	Nominal	“above 50”, “below 30”	one hot encoding	0,1
Qualification of Physician	Nominal	"Dr. med.", "Dipl.-Med.", "Prof. Dr. med.", "Dr.", "MR Dr. med.", "OMR Dr. med.", "Dr.h.c.", "Dr. Dr. med.", "Prof. Dr. Dr. med.", "Dr. med. M.Sc.", "Dr. med. habil.", "Dipl.-Psych.", "Dr. rer.nat.", "Dr. phil.nat.", "Priv.-Doz. Dr. med.", "MD Dr. med.", "Dr. Dr.", "Prof.", "Dr. Dr. phil.", "Dr. phil.", "Priv.-Doz. Dr. Dr."	one hot encoding - All physicians with "Dr." degree as highly qualified and without "Dr." as not highly qualified.	0,1
Insurance type of patient	Nominal	“Private”, “Statutory”	one hot encoding	0,1

Table 5: Categorical attributes and techniques used

Before we proceed to model selection, it is an important observation to note that the “ratings” of the physician (dependent variable) are assessed on a scale of 1-6, 1 being the best and 6 worst, but a mathematical model assumes a higher value as a better grade. To solve this problem, we include an extra step of converting the grades to percentages. The below table demonstrates the range and converted percentage value.

Rating	Grade as Percentage
1	100
(1 - 1.5]	90
(1.5 - 2.5]	80
(2.5 - 3.0]	70
(3.0 - 3.5]	60
(3.5 - 4.0]	50
(4.0 - 4.5]	40
(4.5 - 5.0]	30
(5.0 - 5.5]	20
(5.5 - 6]	10

Table 6: Rating range and converted percentage value

In the next step, we define our dependent and independent variables in our linear model. Since we have multiple independent variables, the type of model suitable for our data is the multiple linear regression model (University, 2018).

This is given by:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \quad \text{for } i = 1, 2, 3, \dots n.$$

Where, y_i is the “grade received by the physician” and x_i is the other “attributes” of the physician. Our data is fit to this model. Now, as a final step in our Regression analysis, we look for the regression coefficients and p-values to interpret the results.

In our study, we form word cloud to understand what patients frequently post about their physicians. We have considered in total 468558 text comments on the General physicians. Word Cloud is a tool which visually represents the existing word frequency and its value. It is used to get an instant insight into the most important terms in the text. In a word cloud, words are shown in varying font size and color depending on how often they appear in the text. The best thing about the word cloud is that it can be generated very quickly. Without a word cloud, it would be a time-consuming job to find the frequency of occurrences of words and to get the context of the text (PluralSight, 2004).

Steps followed to form word cloud:

1. Data collection – The review data (text) from Jameda for the defined sample space of our study i.e., the text reviews by patients on the General Physicians are considered.
2. Data pre-processing – We clean the text by removing the noise and change the text to lower case. The noise here refers to the special characters and emojis used in the text. Converting text to lower case is crucial to match the words and count the frequencies. Furthermore, we remove the words which are not relevant that may hinder our further analysis.
3. Word cloud formation – We check the frequency of occurrence of each word in the text. Finally, we develop the word cloud from the 65 most frequently used words.

3.6 Interpretation of results

To interpret the results of our Hypothesis testing, we first consider a significance level (alpha). Usually, alpha varies from 0 to 1. This signifies the robustness of the null hypothesis. Meaning, smaller the value more robust the model is. In our study, we fix the alpha value to 0.01. The Hypothesis test is done in the previous section, returns a value called p (or p-value). Based on this quantity, we determine to either reject or fail to reject the Null Hypothesis, H_0 (Brownlee, 2018). Now, we compare the p-value to the alpha.

- If p-value > 0.01, we accept the H_0 (implies result is not significant).
- If p-value <= 0.01, we reject the H_0 (implies significant result).

Similarly, in the case of Regression analysis, from the results of our fitted model, we check the p-values of the (independent) variables and compare them with the alpha (Ogee, et al., 2013).

- If p-value > 0.01, then the independent variable is not statistically significant
- If p-value <= 0.01, then the independent variable is statistically significant

The coefficient of the significant independent variables represents the average change in “grade” for one unit of change in the significant independent “attribute”. (Ogee, et al., 2013)

4. RESULTS AND DISCUSSIONS

This section describes the results obtained through the methodologies described in the previous section. For our study and evaluation purposes, we have considered a total count of 345426 graded and registered physicians along with 5172248 reviews which are available in the German PRW - Jameda for the period between 2008 to 2020. To analyze and effectively derive accurate results, we have filtered the available data by defining a suitable sample space and later continued to work on the selected sample data. In our case, the sample space chosen was “General physicians” as it was found that among the “*different Practitioner types*”, the General physician category had the highest number of registered physicians. Hence, we chose the data related to General physicians (number of physicians and the review count) as our sample space.

In our research paper, we have performed statistical data analysis for the selected data sample space by comparing the overall rating received by a physician based on physician attributes and patient characteristics. We begin our analysis by comparing the overall rating received by a physician based on physician attributes and patient characteristics. A hypothesis test and regression analysis are carried out on the compared variables (mentioned in the previous section) where a comparison is made between their respective variables to identify if there exists a significant difference concerning overall ratings received by a physician or overall ratings given by a patient. Finally, we compute the results, i.e., to accept or reject the null hypothesis (H_0) and check if independent variables are statistically significant or not by comparing the p-value to alpha ($p\text{-value} > 0.01$ or $p\text{-value} \leq 0.01$). These results are highlighted in the respective result sections of the attributes considered.

4.1 Physician Popularity

Physician popularity is one of the key attributes related to the physician which we have considered for our study. In simple terms, it highlights how often the physician is preferred and recommended by patients. The following factors define physician popularity:

- Recommendations received from other Physicians
- Recommendations received from other patients
- Physician review count
- Likes received by a physician
- Number of views received on a physician profile

From the below scatter plots, it is visible that as the value of a physician’s popularity factors increases, the rating gets better. Plots are drawn to interpret each factor. This result is also backed by the hypothesis test in which the values are statistically significant ($p < 0.01$).

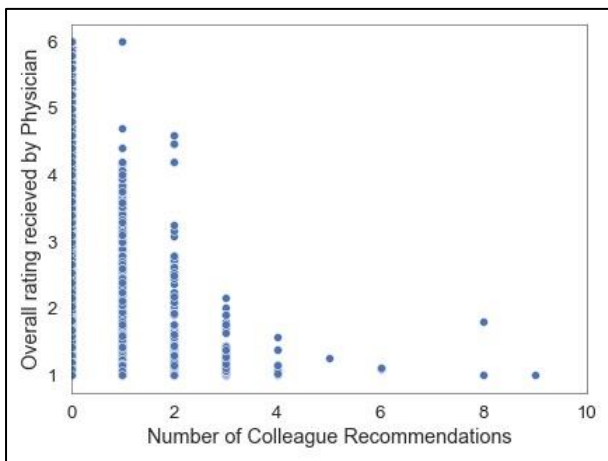


Figure 7: Colleague Recommendations

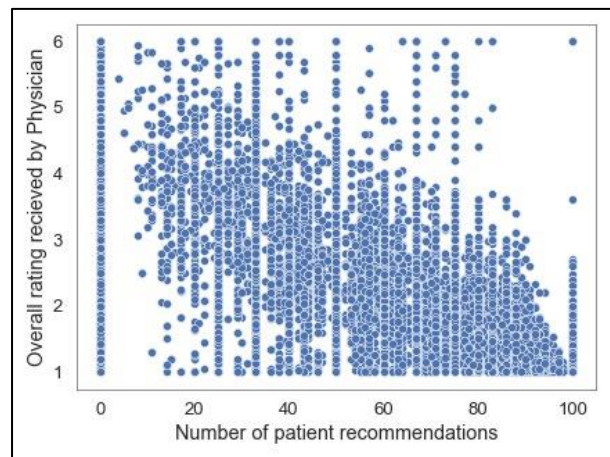


Figure 8: Patient Recommendations

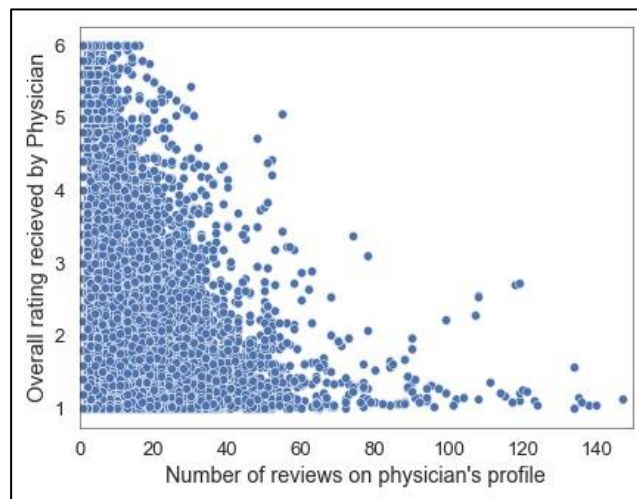


Figure 9: Review Count

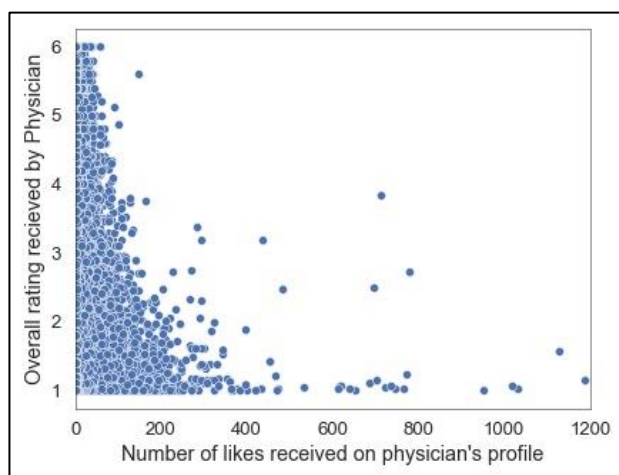


Figure 10: Number of Likes received

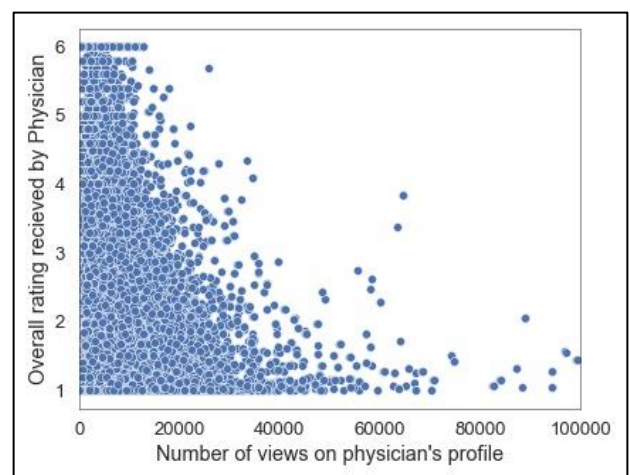


Figure 11: Number of Views

4.2 Physician Gender

The database considered for our study indicates that from the number of registered general physicians, 55.47% are male and 44.53% are female physicians. Based on our findings, it highlights that female physicians receive better ratings than male physicians. The number is small but statistically significant ($p < 0.01$). The below graph gives an overview of the distribution.

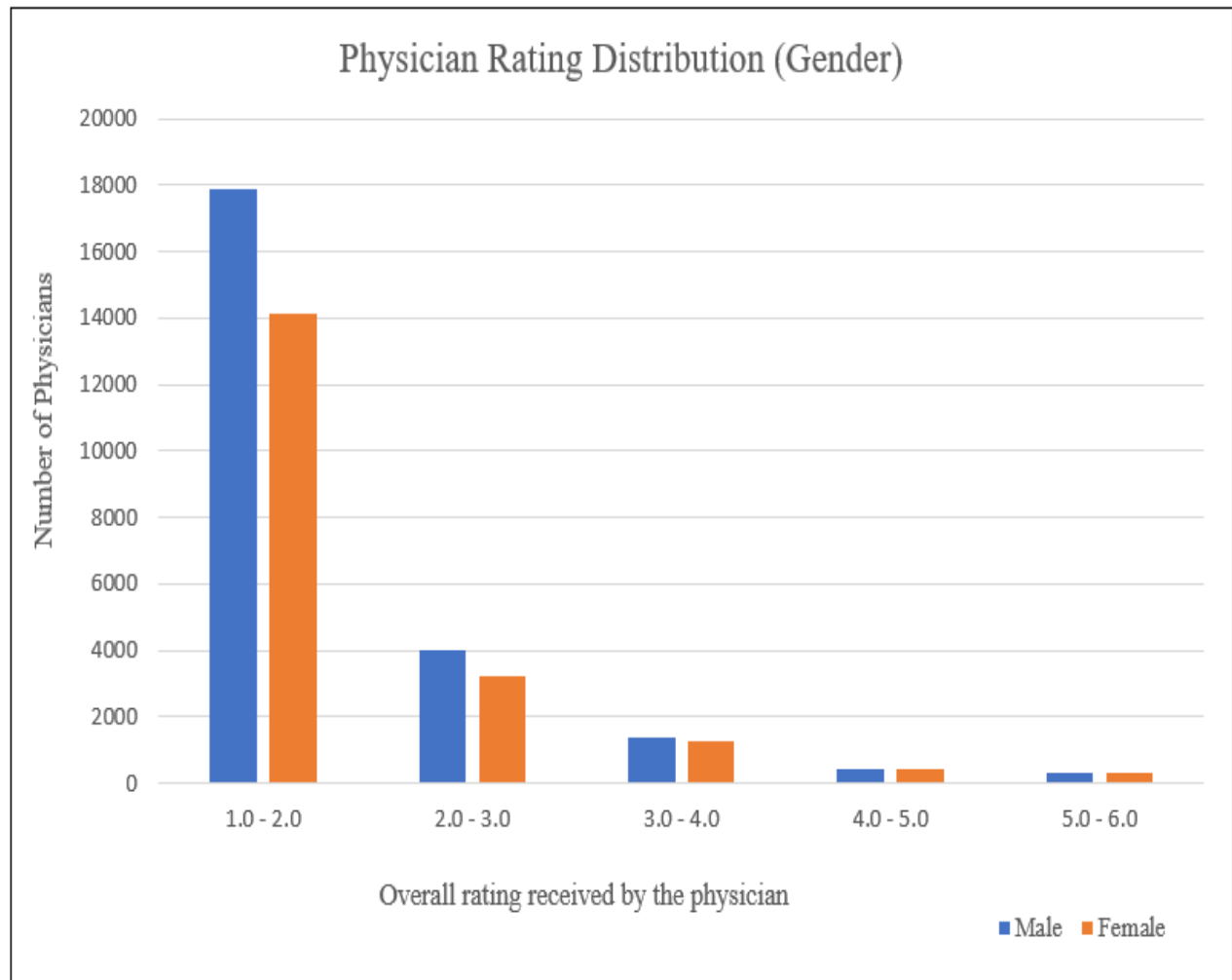


Figure 12: Physician Gender

4.3 Physician Qualification

Physician qualification is one of the attributes which we have considered to answer our research question. Physician qualification defines the highest degree achieved by a particular physician and here we analyze how patients incline towards a particular physician based on their qualification. Our findings tell that 92.16% of the physicians qualify with a “Dr.” degree, while the rest (7.84%) have a qualification with M.A., M.Sc., and physiotherapy. The result from the hypothesis test indicates that physicians with high qualifications receive better grades and the values are statistically significant ($p < 0.01$). The below graph gives an overview of the distribution.

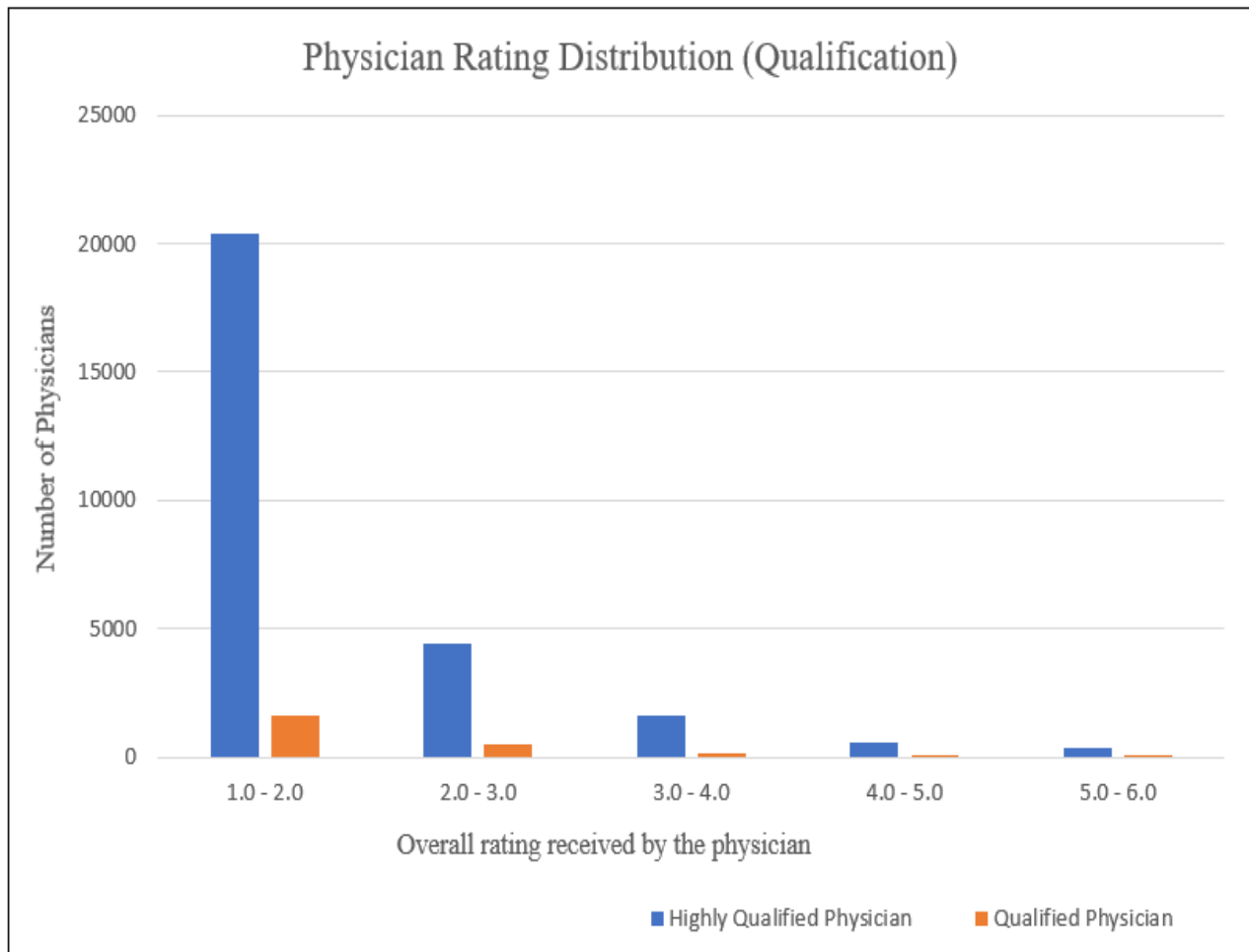


Figure 13: Physician Qualification

4.4 Physician Services

Physician services are referred to as the services offered by the physician, i.e., if the physician is available for consultation at single or multiple locations. Our findings show that 94.96% of the physicians gave consultation at only one location and the rest 5.04% of the physicians gave consultation at multiple locations. When we compared the grade distribution of both the locations, to check if the physicians who provide consultation at multiple locations receive better grades, it was observed that the values were not statistically significant ($p = 0.645$). This infers that Physician Services is not a key factor in influencing patient satisfaction. The below graph gives an overview of the distribution.

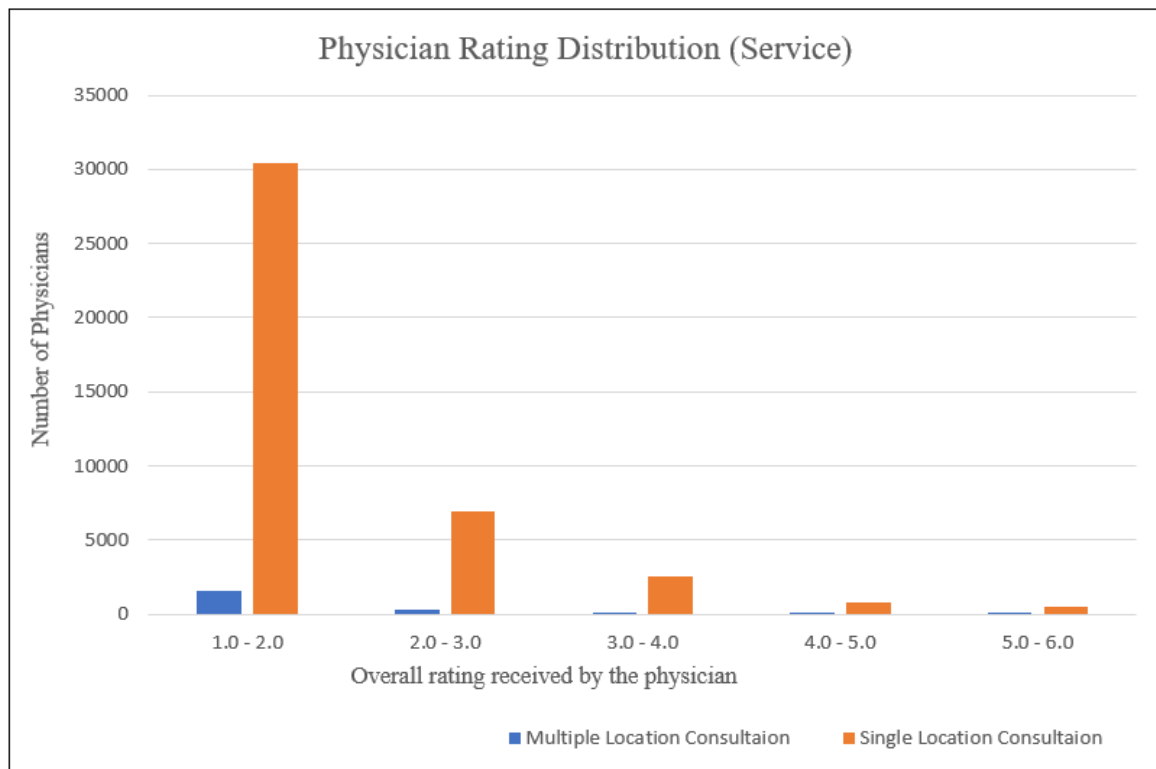


Figure 14: Physician Services

4.5 Type of Insurance covered under a Physician

For our study, we have considered “Type of Insurance “as one of the attributes to see if this factor influences patient satisfaction. During our study, it is seen that every physician has different types of insurance covered under them. As per the data available, insurance type is categorized as follows:

- Private Insurance
- Statutory Insurance
- Type not available

In the data considered for our research, 41.04% of the patients are covered under statutory insurance and just 7.08% under private health insurance. Also, about 51.88% of patients have not disclosed their insurance details on the website. The below graph gives an overview of the distribution of patient insurance type. The result from the hypothesis test indicates that patients covered by private health insurance give more favorable evaluations than patients covered by statutory health insurance. The values are statistically significant ($p < 0.01$). The below graph gives an overview of the distribution.

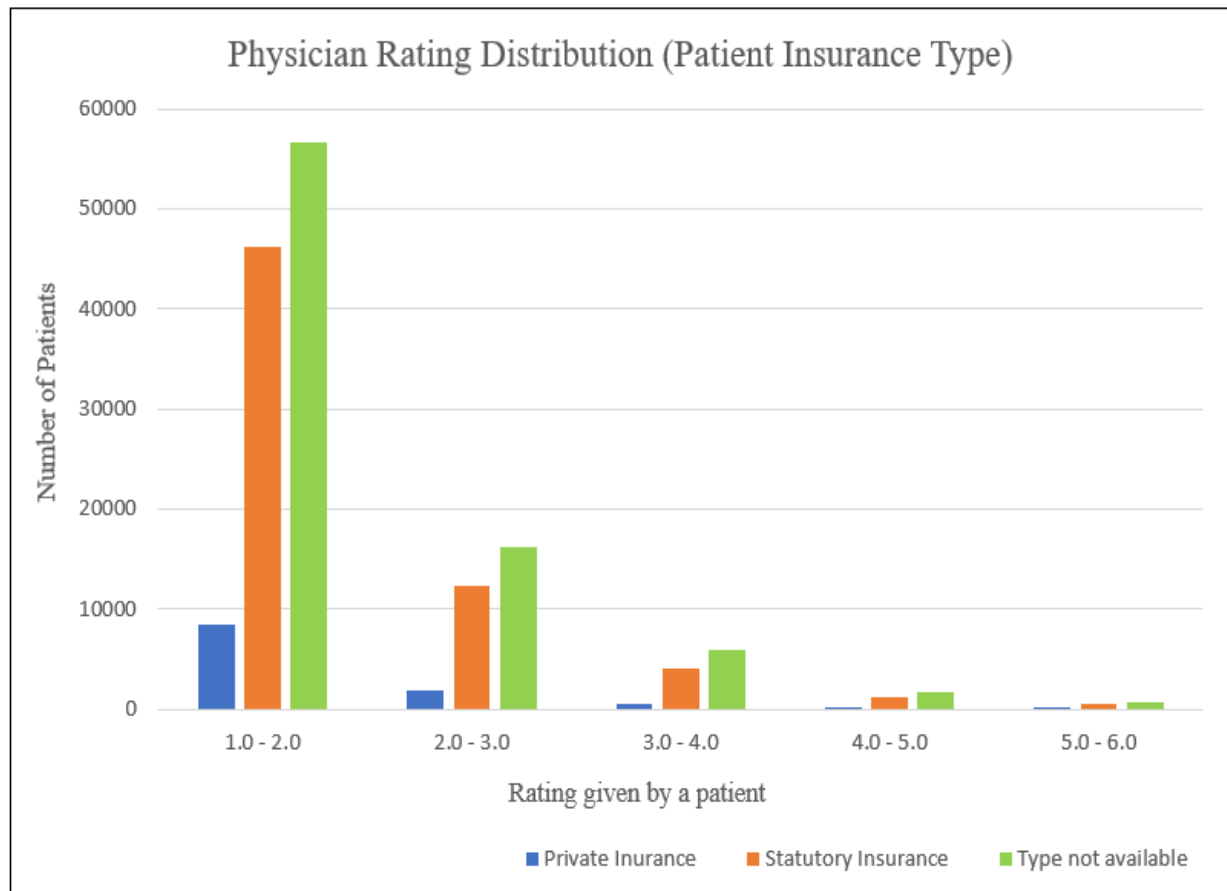


Figure 15: Type of Insurance Covered under a Physician

4.6 Patient Age Group

In our study it is observed that the attribute related to patients age is divided into the following categories:

- Patients' age below 30
- Patients' age above 50
- Patients' age between 30 and 50
- In some cases, the patient has not disclosed their age

Our findings indicate that majority of the patients have not disclosed their age and it accounts for 46.92% of patients, and on the other side patients with the age group above 50 are more (21%) when compared to the age group of patients below 30 (9%). It must be noted that we have not considered the age group of patients who fall between 30-50 years. The result from the hypothesis test indicates that the older patient population gives a better grade than the younger population. The values are statistically significant ($p < 0.01$). The below graph gives an overview of the distribution.

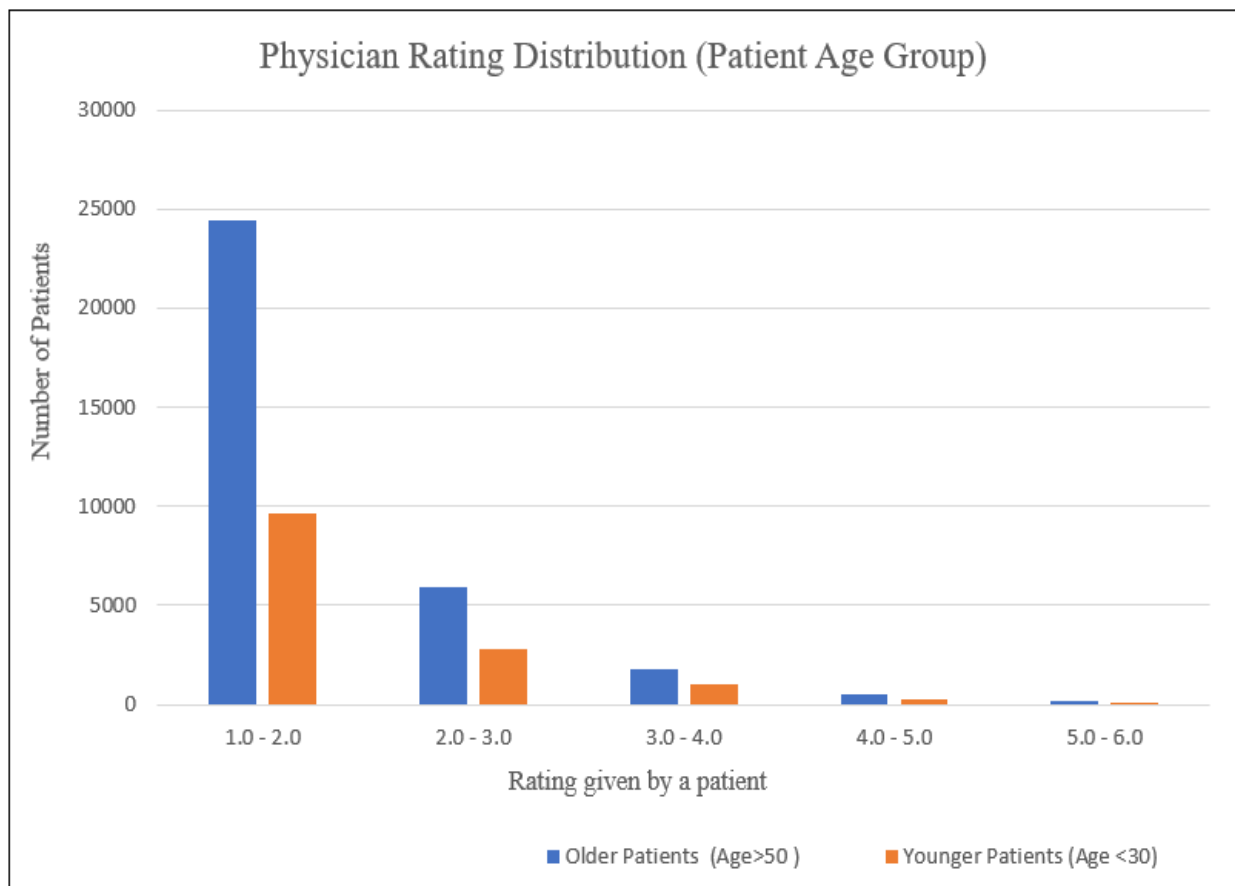


Figure 16: Patient Age Group

To summarize, our study is based on the analysis of 815935 physicians who are registered in German PRW, Jameda from 2008 to 2020. Out of 815925 registered physicians, only 345426 physicians have been graded which accounts for 42% of the total data set. We have considered General Physicians as sample space from this 42% of the data set, which narrows down to 43405 General physicians and 468558 reviews. As shown in Table 7, almost half of the rated physicians were female (44.5%). The majority of the rating patients belonged to the older age group (>50, 21%) and covered Private Health Insurance (7%). The patients preferred also preferred highly qualified physicians (92%).

An analysis was performed to ascertain whether differences in the rating of a physician, regarding both the physician (gender, service, and qualification) and the patient characteristics (age and health insurance) could be determined. The results are displayed in Table 8.

Number and distribution of ratings on Jameda (gender, age, insurance,service and Qualification)				
Characteristics		Absolute	%	%, cumulative
Physician Gender	Female	19330	55.47	55.47
	Male	24075	44.53	100
	Total	43405	100	
Patient Age	< 30	13886	8.88	8.88
	30-50	36272	23.2	32.08
	> 50	32818	21	53.08
	Not available	73334	46.92	100
	Total	156310	100	
Pateint Insurance Type	Statutory Health Insurance	64146	41.04	41.04
	Private Health Insurance	11072	7.08	48.12
	Not available	81092	51.88	100
	Total	156310	100	
Physician Service	Multiple location consultation	2186	5.04	5.04
	Single location consultation	41219	94.96	100
	Total	43405	100	
Physician Qualification	Highly Qualified	27380	92.15752272	92.15752272
	Qualified	2330	7.84247728	100
	Total	29710		

Table 7: Number and distribution of ratings on Jameda (Gender, age, insurance, service and Qualification).

Ratings differences regarding physician and patient characteristics						
Characteristics		N	>Median	<=Median	Percentage below median(%)	P value
Physician Gender	Female	19330	9454	9876	51.09156751	< 0.01
	Male	24075	12228	11847	49.20872274	
	Total	43405				
Patient Age						< 0.01
	< 30	13886	5971	7915	56.99985597	
	> 50	32818	15900	16918	51.55097812	
	Total	46704				
Pateint Insurance Type						< 0.01
	Statutory Health Insurance	64146	31488	32658	50.91198204	
	Private Health Insurance	11072	4896	6176	55.78034682	
	Total	75218				
Physician Service						0.6456
	Multiple location consultation	2186	2627	3196	54.8857977	
	Single location consultation	41219	67418	83069	55.20011695	
	Total	43405				
Physician Qualification						< 0.01
	Highly Qualified	27380	7893	19487	71.1723886	
	Qualified	2330	608	1722	73.9055794	
	Total	29710				

Table 8: Rating differences regarding physician and patient characteristics

Hypothesis Testing				
Sl. No	Hypothesis	p-value	Result	Interpretation
1	Physicians with high popularity receive better grades based on :			
a	Physicians who have more recommendations from their colleagues	0.0	Reject H_0	It is a Key Factor
b	Physicians who have more recommendations from Patients	0.0	Reject H_0	It is a Key Factor
c	Physicians who have more views on their profile	0.0	Reject H_0	It is a Key Factor
d	Physicians who have more reviews from patients	0.0	Reject H_0	It is a Key Factor
e	Physicians who have more likes on their profile	0.0	Reject H_0	It is a Key Factor
2	Female Physicians have better grades than male Physicians	8.7e-70	Reject H_0	It is a Key Factor
3	Physicians who have higher qualification receive better grades	1.3e-47	Reject H_0	It is a Key Factor
4	Physicians who are available on multiple locations	0.645	Fail to Reject H_0	Not a Key Factor
5	Older patient population give better grade to physicians	1.7e-16	Reject H_0	It is a Key Factor
6	Patients covered by private health insurance give more favourable evaluations than patients covered by statutory health insurance	4.4e-27	Reject H_0	It is a Key Factor

Table 9: Hypothesis Testing results

Regression Model						
Sl. No	Coefficients	Estimate	Std. Error	t Value	p-value	Interpretation
1	Physicians with high popularity receive better grades based on :					
a	Physicians who have more recommendations from their colleagues	0.692527	0.166091	4.1696	0.000031	Statistically significant
b	Physicians who have more recommendations from Patients	0.353033	0.001291	273.3822	0	Statistically significant
c	Physicians who have more views on their profile	0.000012	0.000004	3.3263	0.000881	Statistically significant
d	Physicians who have more reviews from patients	-0.048466	0.004266	-11.3605	0	Statistically significant
e	Physicians who have more likes on their profile	0.004173	0.000997	4.1839	0.000029	Statistically significant
2	Female Physicians have better grades than male Physicians	-0.165921	0.137886	-1.2033	0.228863	Statistically not significant
3	Physicians who have higher qualification receive better grades	0.400677	0.233502	1.716	0.086181	Statistically not significant
4	Physicians who are available on multiple locations	-0.213212	0.339269	-0.6284	0.529716	Statistically not significant
5	Older patient population give better grade to physicians	0.523342	0.144937	3.6108	0.000306	Statistically significant
6	Patients covered by private health insurance give more favourable evaluations than patients covered by statutory health insurance	0.364873	0.18172	2.0079	0.044664	Statistically not significant

Table 10: Regression model results

6. LIMITATIONS

- Our study is limited to a certain set of attributes. Few attributes related to the physician such as age or medical school attended were not available in the given database, which could have helped for further analysis.
- One of the factors to be considered is, we cannot gauge the authenticity of information related to patient review and grading, as it could have been subjected to manipulation, resulting in a deviation from our findings.
- Jameda being a German PRW, the study is only limited to Germany, hence the results may vary when considered for different countries. One more factor to be noted is though the sample space is large and is within Germany, our study is limited to only online users.

7. SCOPE OF FUTURE RESEARCH

- In our study we have shown that there exists a significant relationship between physician attributes and patient's satisfaction, there is always a scope for further research on causal analysis. This would be a very interesting study to find if a good rating on the physician is the cause of physician popularity (number of likes, views, reviews, and recommendations), or the other way round i.e., popularity causes good rating on the physician.
- A cross-sectional study could be conducted on data from other websites (possibly from other countries), to validate our current research results. Choosing a different sample space also could lead to many other interesting conclusions as there could be a different set of attributes.
- In our study, we use only word cloud to understand the text comments by patients on physicians. This only lays a foundation but further research on this topic can be carried out for a better understanding of the data. This is possible by using sentiment analysis techniques on text, aggregated from various sources (from both online and offline sources).

8. REFERENCES

1. Bhandari, P., 2020. *Scribbr*. [Online]
Available at: <https://www.scribbr.com/methodology/population-vs-sample/#:~:text=Samples%20are%20used%20to%20make,every%20member%20of%20the%20population>
[Accessed February 2020].
2. Brownlee, J., 2018. *Machine Learning Mastery*. [Online]
Available at: <https://machinelearningmastery.com/statistical-hypothesis-tests/>
[Accessed February 2020].
3. Burkle, C. M. & Keegan, M. T., 2015. Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. *BMC Health service Research*.
4. Damman, O. C., Hengel, Y. K. v. d., Loon, A. J. M. v. & Rademakers, J., 2010. An International Comparison of Web-based Reporting About Health Care Quality: Content Analysis. *Journal of Medical Internet Research*, 12(2).
5. Davis, J., Burrows, J. . F. & Khallouq, B. B., 2017. Predictors of Patient Satisfaction in Pediatric Oncology. *Journal of Pediatric Oncology Nursing*, 34(6), pp. 435-438.
6. Emmert, M. & Meier, F., 2013. An Analysis of Online Evaluations on a Physician Rating Website: Evidence From a German Public Reporting Instrument. *Journal of Medical Internet Research*, August.15(8).

7. Emmert, M. et al., 2014. What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. *Health Policy*, 118(1), pp. 66-73.
8. Emmert, M. et al., 2012. Public Reporting in Germany: the Content of Physician Rating Websites. *Methods Inf Med*, 51(2), pp. 112-120.
9. Emmert, M., Sander, U. & Pisch, F., 2013. Eight Questions About Physician-Rating Websites: A Systematic Review. *Journal of Medical Internet Research*, 15(2).
10. Faber, M. et al., 2009. Public Reporting in Health Care: How Do Consumers Use Quality-of-Care Information?: A Systematic Review. *Medical Care*, January, 47(1), pp. 1-8.
11. Fischer, S. & Emmert, M., 2015. A review of scientific evidence for public perspectives on online rating websites of healthcare. In: *Challenges And Opportunities in Health Care Management*. s.l.:Springer, pp. 279-290.
12. Fox, S., 2014. *The social life of health information*, Washington, DC: Pew Research Center.
13. Fung, C. H. et al., 2008. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med*, 15 January, 148(2), pp. 111-123.
14. Geraedts, M., Hermeling, P. & de Cruppé, W., 2012. Communicating quality of care information to physicians: A study of eight presentation formats. *Patient Education and Counseling*, 87(3), pp. 375-382.
15. Gill, L. & White, L., 2009. A critical review of patient satisfaction. *Leadership in Health Services*, 22(1), pp. 8-19.
16. Greaves, F. et al., 2012. Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. *BMJ Quality & Safety*, 21(7), pp. 600-5.
17. Hanauer, D. A. et al., 2014. Public Awareness, Perception, and Use of Online Physician Rating Sites. *JAMA*, 311(7), pp. 734-735.
18. Hao, H. & Zhang, K., 2016. The Voice of Chinese Health Consumers: A Text Mining Approach to Web-Based Physician Reviews. *Journal of Medical Internet Research*, 18(5).
19. Healthgrades, 2019. *Healthgrades*. [Online] Available at: <https://www.healthgrades.com/> [Accessed 13 February 2021].
20. Hernandez, D. M. K., 2017. *Wordpress*. [Online] Available at: <https://mkhernandez.wordpress.com/2017/04/29/quant-parametric-and-non-parametric-stats/> [Accessed February 2020].
21. Holliday, A. M., Kachalia, A., Meyer, G. S. & Sequist, T. . D., 2017. Physician and Patient Views on Public Physician Rating Websites: A Cross-Sectional Study. *Journal of General Internal Medicine*, 32(6), p. 626-631.
22. Ittersum, K. V., Pennings, J. M. E., Wansink, B. & Van Trijp, H. C., 2007. The validity of attribute-importance measurement: A review. *Journal of Business Research*, pp. 1177-1190.
23. Jameda, 2007. *Jameda*. [Online] Available at: <https://www.jameda.de/jameda/> [Accessed February 2021].

24. Kang, H., 2013. The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, 64(5), pp. 402-406.
25. Lagu, T., Hannon, N. S., Rothberg, M. B. & Lindenauer, P. K., 2010. Patients' Evaluations of Health Care Providers in the Era of Social Networking: An Analysis of Physician-Rating Websites. *Journal of General Internal Medicine*, pp. 942-946.
26. Lebrun, L. A. et al., 2013. Racial/Ethnic Differences in Clinical Quality Performance Among Health Centers. *The Journal of Ambulatory Care Management*, January/March, 36(1), pp. 24-34.
27. Leisen, B. & Hyman, M. R., 2004. Antecedents and consequences of trust in a service provider: The case of primary care physicians. *Journal of Business Research*, pp. 990-999.
28. Li, J. et al., 2018. Developing Embedded Taxonomy and Mining Patients' Interests From Web-Based Physician Reviews: Mixed-Methods Approach. *Journal of Medical Internet Research*, 20(8).
29. McGlynn, E. et al., 2003. The quality of health care delivered to adults in the United States. *New England Journal of Medicine*, June, 348(26), pp. 2635-2645.
30. McLennan, S., Strech, D. & Kahrass, H., 2018. Why are so few patients rating their physicians on German physician rating websites? A qualitative study. *BMC Health Services Research*, 18(1).
31. Mukamel, D. B. & Mushlin, A. I., 1998. Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports. *Medical Care*, July, 36(7), pp. 945-954.
32. Ogee, A., Ellis, M. & Pammer, C., 2013. *The Minitab Blog*. [Online] Available at: <https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients#:~:text=Regression%20coefficients%20represent%20the%20mean,predictors%20in%20the%20model%20constant.&text=The%20coefficient%20i> [Accessed February 2020].
33. Otani, K. et al., 2009. Patient satisfaction: focusing on "excellent". *J Healthc Manag*, pp. 93-102.
34. PluralSight, 2004. *PluralSight*. [Online] Available at: <https://www.pluralsight.com/guides/natural-language-processing-visualizing-text-data-using-word-cloud> [Accessed February 2020].
35. Porter, M. E. & Guth, C., 2012. *Chancen für das deutsche Gesundheitssystem*. Berlin: Springer/Gabler.
36. Rothenfluh, F. & Schulz, P. J., 2017. Physician Rating Websites: What Aspects Are Important to Identify a Good Doctor, and Are Patients Capable of Assessing Them? A Mixed-Methods Approach Including Physicians' and Health Care Consumers' Perspectives. *Journal of Medical Internet Research*, 19(5).
37. Rothenfluh, F. & Schulz, P. J., 2018. Content, Quality, and Assessment Tools of Physician-Rating Websites in 12 Countries: Quantitative Analysis. *Journal of Medical Internet Research*, 20(6).
38. Schaarschmidt, M. & Ivens, S., 2017. *Dr. Miller or Dr. Smith ? Patients' Intentions to make appointments on Physician Rating Platform*. South Korea, s.n.

39. Terlutter, R., Bidmon, S. & Röttl, J., 2014. Who Uses Physician-Rating Websites? Differences in Sociodemographic Variables, Psychographic Variables, and Health Status of Users and Nonusers of Physician-Rating Websites. *Journal of Medical Internet Research*, 16(3).
40. University, Y., 2018. *Yale Educational institute*. [Online]
Available at: <http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm>
[Accessed February 2020].
41. Verhoef, L. M. et al., 2014. Social Media and Rating Sites as Tools to Understanding Quality of Care: A Scoping Review. *Journal of Medical Internet Research*, February.16(2).
42. Volpi, G. F., 2020. *Medium*. [Online]
Available at: <https://medium.com/swlh/what-are-categorical-variables-and-how-to-treat-in-machine-learning-fd078c09a9af>
[Accessed February 2020].
43. Wilkie, W. L. & Pessemier, E. A., 1973. Issues in Marketing's Use of Multi-Attribute Attitude Models. *Journal of Marketing Research*, pp. 428-441.
44. Zeithaml, V. A., Berry, . L. . L. & A. P., 1993. The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science*, 21(1), pp. 1-12.