Friedrich-Alexander-Universität Erlangen-Nürnberg
Technische Fakultät, Department Informatik

DO MINH CHAU LE

MASTER THESIS

MARKET RESEARCH FOR PERSONAL USE OF QDACITY

Submitted on 24. November. 2021

Supervisor: Dr. Ing. Andreas Kaufmann, Prof. Dr. Dirk Riehle, M.B.A.
Professur für Open-Source-Software
Department Informatik, Technische Fakultät
Friedrich-Alexander University Erlangen-Nürnberg

Versicherung

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

_____

Erlangen, 24.11.2021

License

_____

Erlangen, 24.11.2021

Abstract

Computer-aided qualitative data analysis software (CAQDAS) supports coding, categorizing, thematic analysis and derivation of insights from qualitative data. This category of software is most frequently found in academic settings but not exclusive to those. Thinking of the potential commercial value of such a solution to researchers, industry market research professionals, as well as scholars and students as well as companies and institutions, the Open-Source Research Group has been planning the launch of a cloud-based CAQDAS. This thesis, structured through a list of individual hypotheses, is an effort to figure out the potential market and business value of private use of a specific cloud-based CAQDAS. The thesis combines primary data collection using a survey and secondary sources like market research, reports, statistics. The survey findings serve to confirm or disconfirm the individual hypotheses. The conclusion to each hypothesis combined with insights from competitor analysis and secondary research are translated into business recommendations for the development team, in crafting their business plan for the product. The main findings from the research indicate potential business value for launching QDAcity within the mid-range segment with corresponding pricing and product proposition.

Content

List of Figures

List of Tables

1 Introduction

First emergent in 1980s, computer-aided qualitative data analysis software (CAQDAS) has transformed the way qualitative research is carried out (Costa et al., 2017). Over time, the functionalities of CAQDAS have no longer been limited to just retrieving, coding, categorizing qualitative data but also expanded to including complex analysis qualities using mixed-methods, visualization, varied processing and outputting of data types. Some of the earliest CAQDAS have evolved to sophisticated solutions and claimed leading positions in the market, for example NVivo (earlier as NUD*IST), WinMAX (MAXQDA), ATLAS.ti. Besides these popular names, there are also simple coding software that comes with no fee, i.e. Taguette, QDA miner Lite, RQDA (an extension of R for qualitative data analysis) (Costa et al., 2017) (Freitas et al., 2017). While being less popular than quantitative data analysis tools with much fewer options in the market, CAQDAS is still a very useful tool for researchers and students to deliver their research work. This makes CAQDAS an appealing area for new businesses.

This thesis is therefore an effort to investigate the potential market and business value in personal, private use of QDAcity – a cloud-based computer-aided qualitative data analysis software (CAQDAS) currently in development by Open-Source Research Group at University of Erlangen-Nürnberg (FAU). The contributions of this thesis are as following:

- To figure out if there is a potential market for personal use of QDAcity
- To illustrate and estimate the target markets
- Through multiple hypotheses, to deliver practical business recommendations

Besides, considering limited amount of existing literature on the market or market research of these tools, this paper also serves to explore potential private users' preferences towards offerings and functionalities of CAQDAS. This may provide useful hints and gaps for further research in the future on similar or relevant topic, i.e. market potential of CAQDAS.
The paper continues with Related Work (Chapter 2), Research question breakdowns and operationalization (Chapter 3), then approach in research and data collection, including a structure and reasoning of survey questions (Chapter 4). Next, the paper delivers findings from the survey and secondary research in Chapter 5 and formulates the information into a marketing plan for QDAcity in Chapter 6 - Discussion. Limitations of the research and opportunities for further work are also presented in Chapter 6. Finally, Chapter 7 delivers the overall conclusion to the thesis.

2 Related Work

Existing research work mostly addresses approaches to qualitative data analysis using software tools and fundamental concepts in qualitative data analysis: Silver and Lewins (2014), Woods et al. (2016), O'kane et al. (2019), use cases of CAQDAS, industries and/or with specific purposes: mixed methods, visualization like Prabowo, (2020), Dalkin et al. (2020), Humble (2020), Verdugo-Castro et al. (2019), benefits and challenges of tools towards qualitative data analysis, the future of CAQDAS and a wish list of functionalities in research practice: Evers, 2018, the possibility and ways of using a general-purpose software (i.e. Microsoft Word) for qualitative data analysis: LaPelle (2004). Some others elaborate on the use of specific solutions in this category, e.g. MAXQDA, by Kuckartz (2019), QSR NVivo by Jones (2007) and users' learning to use a CAQDAS by Freitas et al. (2017) and Kalpokaite and Radivojevic (2019). There has also been research work about the benefits of some qualitative research practices such as interrater reliability (agreement) and member checking like those by Gisev et al. (2013), Birt et al. (2016).

The existing literature serves mostly to provide reference for fundamental concepts of qualitative data analysis and basic functionalities of a CAQDAS including coding, categorizing, visualizing, saturation, etc. These concepts are used in this thesis to elaborate on what type of software a CAQDAS is, which functionalities and features it provides, how it facilitates qualitative research and data analysis. However, no existing literature appears to address questions about commercial value and potential market of a CAQDAS. Thus, this chapter will not go on elaborating about key findings and approaches of each of the mentioned literature, because little of them actually serves to provide directly relevant inputs for the topic. Despite addressing the same subject matter, the literature, mostly using literature review presents findings about other aspects of CAQDAS rather than market potential, users' needs and behavior. This thesis, in contrast, focuses on the market, commercial value of CAQDAS and uses empirical data collection to arrive at final conclusions. It facilitates, therefore, not only the creation of a business plan for QDAcity itself but also a foundation for market research, further validation effort and follow-up study in improving or launching a CAQDAS.

The only points in common, as mentioned, are fundamental concepts of qualitative data analysis and use of CAQDAS in qualitative research. The next section will utilize the information in existing literature to operationalize the used terms and concepts with respect to qualitative data analysis and CAQDAS in this thesis.

3 Research Question

There are two different segments of customers for a CAQDAS. One is at institutional or organizational level where purchase of tools is done by a budget holder for use of multiple teams and people. The other segment refers to groups of users who buys out of their pocket or adopts a tool for their personal use, without being sponsored by institutions. Within the scope of this thesis paper, the institutional customers will not be considered and only use of individual end-users will be. In other words, the paper seeks to validate that a CAQDAS is also commercially valuable for private and individual use.

Thus, the research question for this paper is as mentioned:
Is there a potential market for personal use of QDAcity, given its pricing and proposition?

The following hypotheses serve as a step-by-step approach to tackle multiple aspects of the product that allow market size estimations, sales forecasts and business recommendations:

1) Potential market:
a. There are potential users who are willing to pay for personal use of CAQDAS.

2) Product proposition and pricing:
b. There is demand for a lightweight and focused CAQDAS.
c. A majority of students would opt for a free option if they used a CAQDAS.
d. There are potential users who have concerns about storing their data online.
e. Concerns about storing research data online do not prevent most users from adopting a cloud-based CAQDAS.
f. Potential users who are willing to pay prefer a rolling subscription.

3) Functionality preferences:
g. Offline use is an important feature for a CAQDAS.
h. Real-time group coding is an important feature for a CAQDAS.
i. Measure of interrater reliability is an important feature for a CAQDAS.
j. Support for member-checking (permission schemes) is an important feature for a CAQDAS.
k. Transcribing audio/video files is an important feature for a CAQDAS.
l. Coding of audio/video files is an important feature for a CAQDAS.
m. Support for processing multiple file types is an important feature for CAQDAS.
n. Saturation measure is an important feature for a CAQDAS.
o. Professional technical support (e.g. call center) is an important feature for a CAQDAS.
p. Dictionary is an important feature for a CAQDAS.
q. Mobile version or portability is an important feature for a CAQDAS.
r. Autosave, versioning and backup options are important features for a cloud-based CAQDAS.
t. There is demand for integrations between the CAQDAS and some third-party applications.

Operationalization of used terms

Below is the table of working definitions of terms used in this thesis. As mentioned, most of these are fundamental concepts for qualitative data analysis which align also with basic features and functionalities of a CAQDAS.

Used term	**Definition**
CAQDAS	a computer-assisted qualitative data analysis software, either on-premise or cloud-based (Silver & Lewins, 2014) (Kuckartz, 2019)
Qualitative data analysis	coding textual and non-numerical information (Silver & Lewins, 2014)
Coding	labeling and categorizing qualitative data to identify themes and relationships (Silver & Lewins, 2014) (Kuckartz, 2019)
Interrater reliability	a measure of agreement in independent coding among two or more coders (Gisev et al., 2013)
Member checking	a practice to validate accuracy of categorization of collected data with respondents or participants (Birt et al., 2016)
Saturation	the point where no new category can be extracted from the remaining data (Silver & Lewins, 2014) (Glaser & Strauss, 1967)
Lightweight and focused CAQDAS	a solution that caters the basic needs of users in qualitative data analysis, including imports and exports of most used file types (e.g. text), coding, categorizing, summary of coded texts (with or without basic graphs). Examples: Taguette, QDAminer Lite.
Feature-rich CAQDAS	a solution that caters a more extensive need in working with qualitative data, including processing of multiple different file types, multi-method analysis, sophisticated visualization, collaboration measures, or even up to machine learning and automation features. Examples for this group of CAQDAS are MaxQDA, Atlas.ti, NVivo.

Table 1. Definitions of used terms in the thesis

4 Research Method

The market research is done combining the following methods and sources of data:

- Secondary research for understanding market, trends and general situation of competition. This includes keyword search, backward referencing of news, websites of competing products, statistics and articles.
- Primary research using survey for determining segment size, customer profile, habits, use patterns, key features and any potential for integration.

4.1 Question content and purposes

Each question included in the survey was selected to elicit certain insights from the potential users. They can be grouped as follows:

4.1.1 Hypothesis addressing questions

A majority of the questions aim to investigate potential users' needs, behavior and preferences, contributing to confirming or denying the individual hypotheses.

4.1.2 Hidden insight questions

Some questions go deeper in exploring the options not listed in the survey itself, for example those asking respondents to list features they need in a CAQDAS or an appropriate pricing. Here potential users can freely provide their opinions, thoughts which can be interesting for a marketing plan of QDAcity.

Questions about frequency of use and habits allow for a better understanding of how users use the respective product which can affect their purchase decisions. This insight is helpful for pricing and plan offers. These also identify if a sustainable and considerable demand is actually there or respondents just fill in the other questions regarding preferences and purchase decisions arbitrarily.

4.1.3 Profile questions

These help explore any trends specific for any sub group of the target customers. Also, they give details of people who answer the survey and whether reaching out to certain groups of identified customers is required to maintain the reliability of the survey results.

Specific questions are presented in the Survey Structure.

4.2 Question formats

4.2.1 Multiple choice – Single answer

These are used to classify users or respondents, either they are interested in the type of product or not, how often they use the product or have the need to. The answer options are mutually exclusive and thus put each respondent in a specific class based on their answers.

4.2.2 Multiple choice – Multiple answers

The purpose for using this type of questions is to get to know the preferences and specific needs of the users regarding how they use the product. This may include several predefined items which are not mutually exclusive. These predefined options are the most popular and easiest to think of, considering what the competitors offer, existing research on users' behavior and use of the same product type. The list of options is nonetheless not exhaustive. It has also a free-text option, i.e. "Others:" where respondents can fill in any other input which is not already mentioned.

4.2.3 Likert scale

This type of questions is for checking how important certain needs or preferences are to users. Likert scale questions ask respondents to rank different items in terms of importance from a scale of 5 or 7 with an option to be neutral (indifferent). It can be interesting that some users may have certain needs but they value them not equally and do not feel it is very critical to have some of them covered in the respective product (already have that with another product / tool which works just well). Such insights allow decisions of which features to prioritize, to include in core product or upgraded version, or not to consider at all.

4.2.4 Open ended

These are used to explore the unknown. Respondents can freely provide their opinions, ideas without being probed by given options. These are helpful to gain insights into customers' needs, preferences, usability problems and so on, that we could not identify initially.

Some respondents may know the software but are not (yet) interested in purchasing or adopting a product of same kind. In this case, it would be helpful to ask for their thoughts and opinions on pricing and product features which can give hints on what to improve or what could attract them to purchase. If those preferences are feasible and can be part of the upcoming launch of product, there is more likelihood that the product will attract a larger market.

4.3 Survey structure

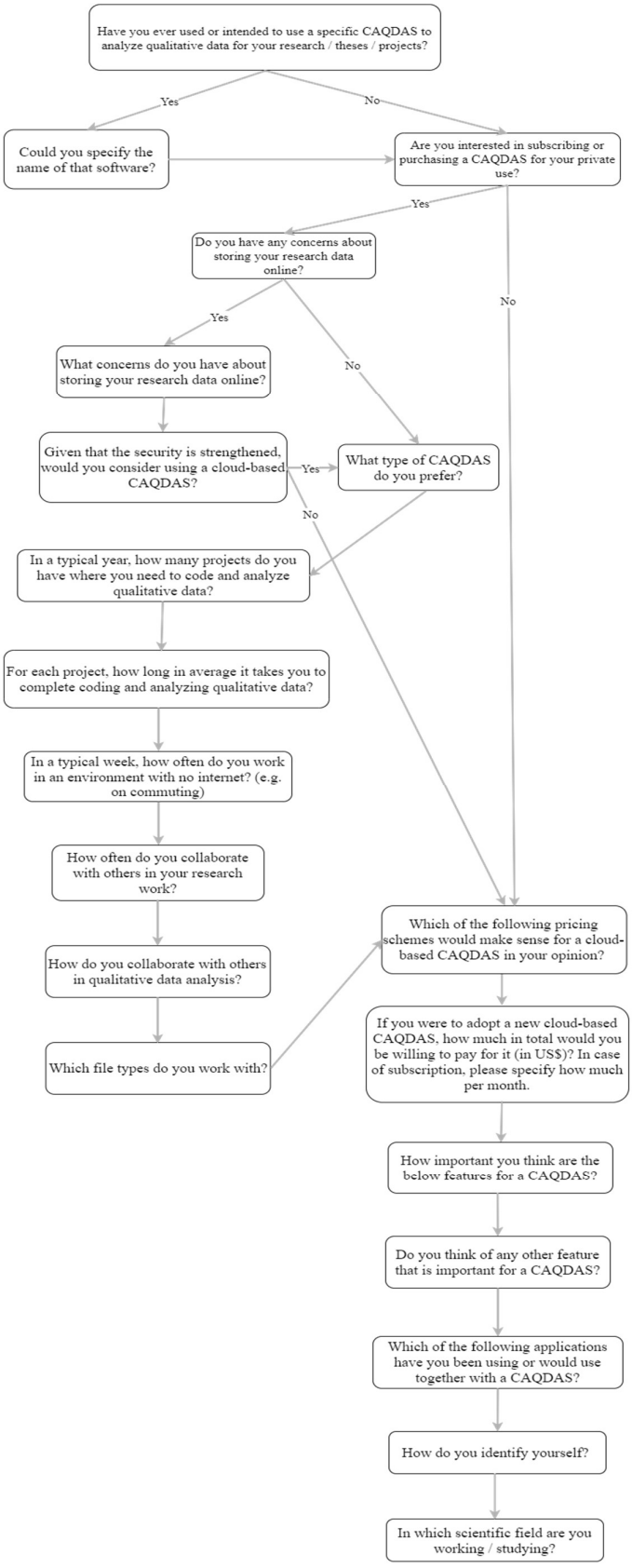

Figure 1. Survey structure

Figure 1 shows the flow of questions in the survey. The questions are not subsequent but rather pre-conditioned and event-based.

In case respondents have used a CAQDAS before, they will be asked which specific solution or brand name it is. If they have security concerns in a cloud-based solution, they will be asked which particular concerns they have and whether or not they would still buy if security were strengthened. If no concerns, they will be asked about their preferred CAQDAS type.

In case the respondents are not interested in purchasing a CAQDAS or refrained from using one due to security concerns, they will not be asked about their research activities and habits, because they are very much likely not to opt for a cloud-based solution like QDAcity. However, these are still probed for opinions regarding pricing and features since such insights are useful for a CAQDAS no matter how it is hosted.

The reason for doing this is to ensure that respondents need not to waste their time on questions not directly relevant for them, and thus make the survey shorter, simpler and respondent-friendly. In addition, with some embedded logic, analysis can also be easier. For example, when aggregating data about research behavior and habits, we take into consideration only people who have interest and are not deterred by security concerns in adopting a cloud-based CAQDAS. This allows a breakdown of data directly without an additional filter or aggregation of multiple attributes.

4.4 Data collection process

The following ways were used to seek responses for the survey:

- Reaching out to Researchers, PhD, Bachelor and Master students in Facebook groups, i.e. International Students in Germany, France, Netherlands, Finland, United Kingdom, Canada, Australia, User Experience Researchers, Vietnamese PhD candidates, Vietnamese students in Germany, International Scholars, Survey exchange and scholarship information groups. A detailed list of these Facebook groups is presented in Appendix A.
- Searched for email addresses of university researcher, research assistants on faculty websites of universities, including FAU, mostly in Germany, then sent bulk emails to ask for support to fill in the survey. List of contacted universities can be found in Appendix A.
- Asked for support from a program or faculty coordinator, in this case the Faculty of Engineering, to distribute the survey to all students of the faculty.
- Shared the survey also in FAU Facebook groups, i.e. for Medical Engineering students, and groups of other purposes like house searching, flea market where many students at FAU gather.

The survey was expected to reach about 2,000 people in total, researchers and researcher assistants or students in universities in Germany, United Kingdom and the Netherlands via direct emails. For free sharing and spontaneous response on social media platforms, it is hard to know exactly how many people were reached. Nonetheless, each social media group where the survey was shared has at least 1,000 members (except for FAU International Master Computational & Medical Engineering with only 702 members of the faculty itself).

Out of all these 141 accepted to fill in the survey and gave a response, among which 102 are complete responses.

5 Research Results

5.1 Survey results

5.1.1 Data cleaning and preparation

A workbook is provided together with this thesis paper to allow reference to raw and aggregated data. The sheet "All" in this workbook records all 141 responses, both complete and incomplete. Meanwhile, the sheet "Complete" records only 102 complete responses. Incomplete responses will not be used for data aggregations because they end before the question about respondents' occupations. Since aggregations are done based mostly on user groups (respondents' occupations), it is of no use to include incomplete responses. However, as incomplete responses can serve to provide additional useful insights, especially for open-ended questions about used file types and preferred features, they are still included for analysis.

Before the data are analyzed, some adjustments have been made to ensure consistency and plausibility. Specific amendments to the survey results are listed as follows:

5.1.1.1 Related questions with inconsistent answers

Some inconsistent answers have been identified for the following two survey questions:

- Have you ever used or intended to use a specific CAQDAS to analyze qualitative data for your research / theses / projects?
- Could you specify the name of that software?

Around 7 percent of the complete responses (7 out of 102) claim to have used a CAQDAS before. However, when asked to specify the name of the software, they list instead a quantitative analysis solution and/or programing language or mentioned explicitly that there is none they are aware of (Table 2). For this portion of responses, the answer to the first question ("Have you ever used…") is changed to "No" to ensure consistency. In the corresponding data sheets "Complete", the adjusted answers are highlighted in yellow.

Ever used	Name	Changed – Ever used
Yes	SPSS	No
Yes	R, stata, python	No
Yes	SQL (SQLite etc)	No
Yes	R, pandas, numpy, scilearn	No
Yes	I am still searching	No
Yes	No	No
Yes	No software that I am aware of.	No

Table 2. Conflicting responses to related questions

5.1.1.2 Invalid, irrelevant responses

A whole response is removed when the answers to multiple questions are irrelevant or not meaningful at all for the analysis or recommendations of this thesis paper. Amongst the 102 received complete responses, there is one that continuously claims a

lack of explanation of terms and concepts while a glossary is already provided in the first page of survey.

Some answers in this response are:
“I don’t know what caqdas is, maybe explain it with less jargon if you are going to send this survey to people who have not even heard of this stuff.”
“none. I already said I don't know what caqdas is….”

Due to the lack in meaningful inputs and understanding of the respondent of relevant terms and concepts, this whole response would be skipped for consideration.
After removal of this response, the total number of complete responses is 101.

5.1.1.3 Standardizing pricing inputs

This section is dedicated for responses about pricing. Specifically, regarding the question “If you were to adopt a new cloud-based CAQDAS, how much in total would you be willing to pay for it (in US$)? In case of subscription, please specify how much per month.”, all the answers are given in free text and thus not standardized. For the purpose of aggregating, they are standardized as follows in the workbook:

1) A duplicated sheet of “All” named “Price mapping” is created, which includes all responses, complete or not. Since several respondents stopped the survey right after this question, their answers can still help to provide insights. A creation of a new sheet is to ensure the data are isolated for pricing-related changes only. For other hypotheses or aggregations using other attributes, the original entries in sheets “All” and “Complete” are used.

2) In this new sheet, a new column named “Price given” is created. This column illustrates the manner in which the respondents choose to quantify their willingness to pay, either per month or as a lump sum, based on their answers.

3) Another new column named “Max price” is added. If the respondent gives an individual number, the same number is entered in this column. Otherwise, if a range is given, the upper limit is chosen. It shows the maximum the respondents are willing to pay for a CAQDAS. For values such as “less than x”, a value of 0.01 less than x is used (e.g. 200 and 199.99). 0.01 may not be trivial to account for, however, research has proved that such a difference is not only visible but also attractive to customers, as they feel like they are getting a bargain (Robert, 1991). Thus, a customer who is willing to pay for 199.99 may not be willing to buy the same thing at the price of 200 USD.

4) In addition, there are three responses which provide both a lump sum and price per month:

   “In total, it should not cost more than $100. In case of subscription, it should not cost more than $10 a month. I give these number based on the current prices of Microsoft 365, survey monkey (which is quite high), and other cloud based services.”
   “20 p. mo. ; 200 one time”
   These entries are duplicated to account for the correct frequency each type is mentioned. In general, the total entries after duplication amount up to 144 from

141 originally. This adjustment is limited only to pricing-related data and thus is made only to the duplicated sheet of survey data ("Price mapping").

5.1.2 Data interpretation

The insights gained from the survey responses provide either confirmation or disconfirmation to the research hypotheses. For each hypothesis, there will be a conclusion if a feature, an option or direction is either taken forward in this thesis or not. Eventually we will connect these dots to deliver an overall conclusion and recommendations with respect to product, pricing, placement and promotion (4P Strategy) of QDAcity.

1) Potential market:

a. There are potential users who are willing to pay for personal use of CAQDAS.

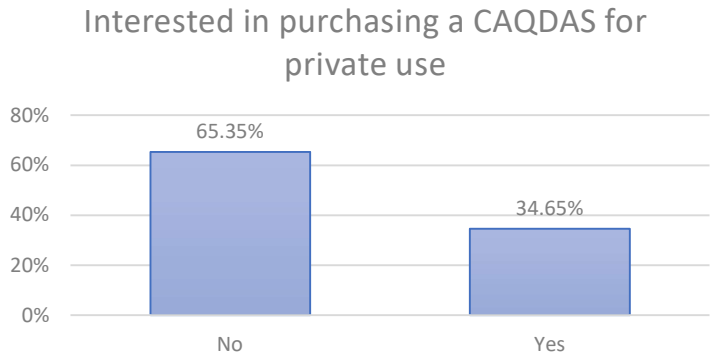

Figure 2. Respondents' interest in purchasing a CAQDAS for private use

As shown in Figure 2, there are respondents (34.65 percent out of all) known also as potential users who are willing to buy a CAQDAS for their own use. It means there could be a potential market for QDAcity.

2) Product proposition and pricing:

b. There is demand for a lightweight and focused CAQDAS.

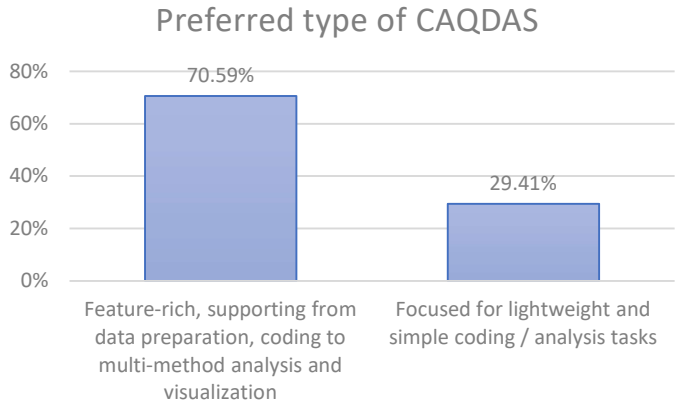

Figure 3. Respondents' preferred type of CAQDAS

From Figure 3, it can be seen that out of people who are:

- interested and
- have no security concerns or will still buy if security is strengthened

approximately 29 percent opt for a focused and lightweight solution. There is thus a considerable difference in demand between the two types of software.

c. A majority of students would opt for a free option if they used a CAQDAS.

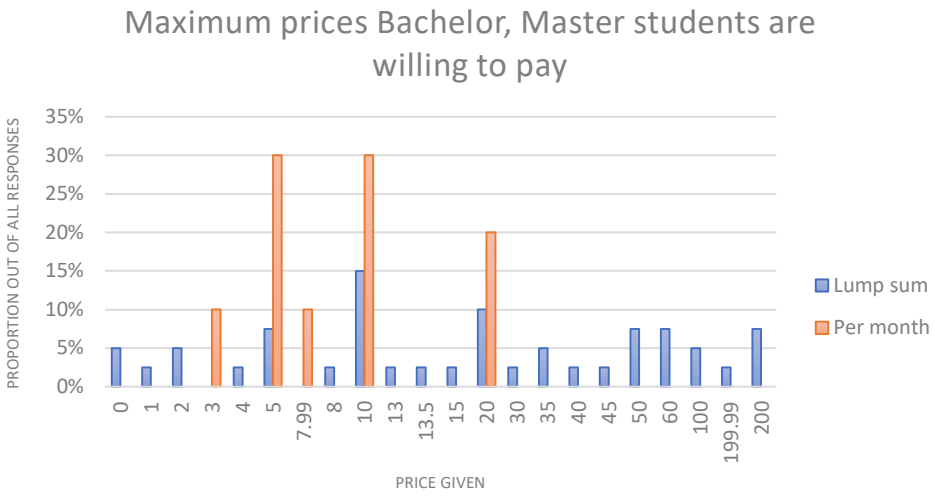

Figure 4. Maximum prices Bachelor, Master students are willing to pay

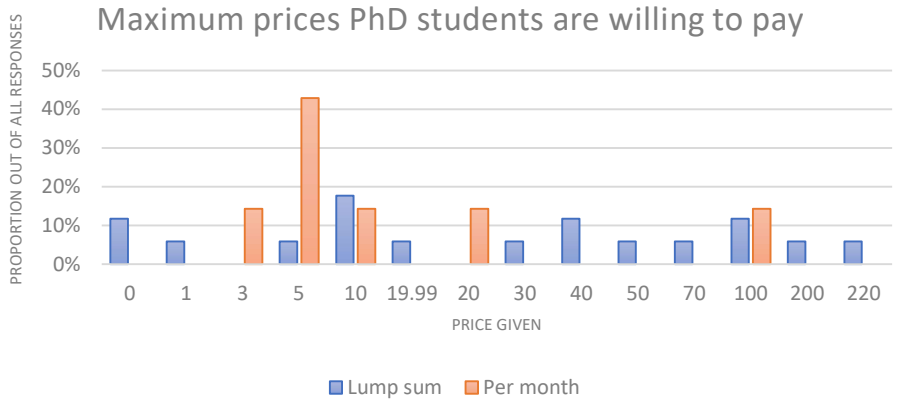

Figure 5. Maximum prices PhD students are willing to pay

It can be seen from Figure 4 and 5 that a majority of students who are interested in purchasing a cloud-based CAQDAS are willing to pay at least something for it: only 5 percent for Bachelor, Master students and 11.76 percent for PhD students, counted only among given lump sum values, give a 0 for this question. If counted among all given values, the frequency of 0 would be even lower: 5.88 percent for Bachelor, Master students and 8.33 percent for PhD students.

There could be arguments that these respondents might have been afraid of being judged for putting 0. However, given the survey is completely anonymous and all groups of potential users were approached similarly, there should not be a difference between them in terms of willingness to put 0 as a price. As students, it is actually even more comprehensible for some to claim that they prefer a free software, because of their limited financial ability.

As shown in Figure 4 and 5, the frequency is more spread across different price levels when it is a lump sum than when it is a monthly subscription. The price of 10 USD one-time is the most frequently listed by both Bachelor, Master and PhD students, though not significantly more compared to remaining values. For cost per month, most of the votes go to 5 USD and 10 USD.

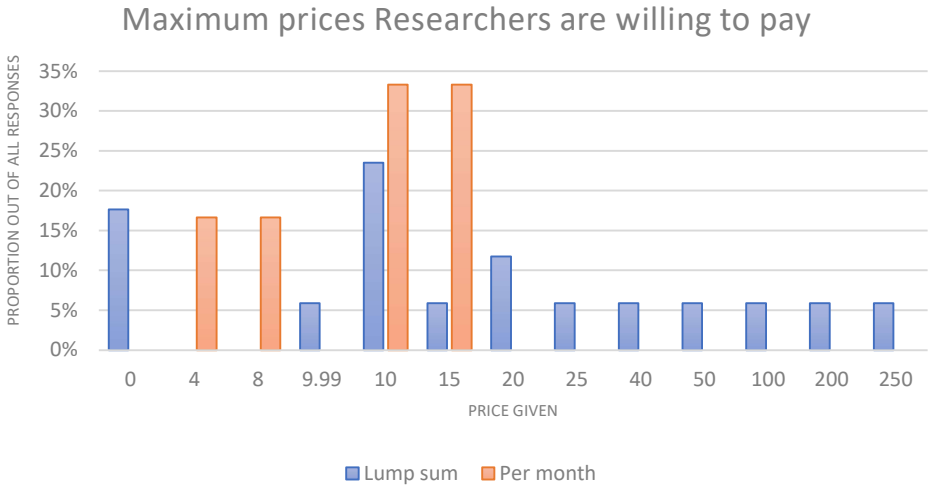

Figure 6. Maximum price Researchers are willing to pay

Meanwhile for researchers, statistics show that around 17.65 percent of respondents who give a lump sum are not willing to pay anything for a CAQDAS. This translates to around 13 percent among all Researchers who give a value.

As shown in Figure 6, 10 USD one-time is the most frequently listed price by Researchers. For cost per month, 10 and 15 USD are the most picked.

It can be seen from most figures for both students and researchers that the one-time cost, which is supposed to be much higher, is actually lower or equal to the monthly cost. This is as per common sense is not plausible. However, it shows a tendency among all respondents who answered this question that their perception of "one-time" is rather relative. The service may not need to last a life-time but instead long enough for their purpose within a specific period, i.e. coding and analysis of research work, a service they can take and drop anytime when they feel like.

d. There are potential users who have concerns about storing their research data online.

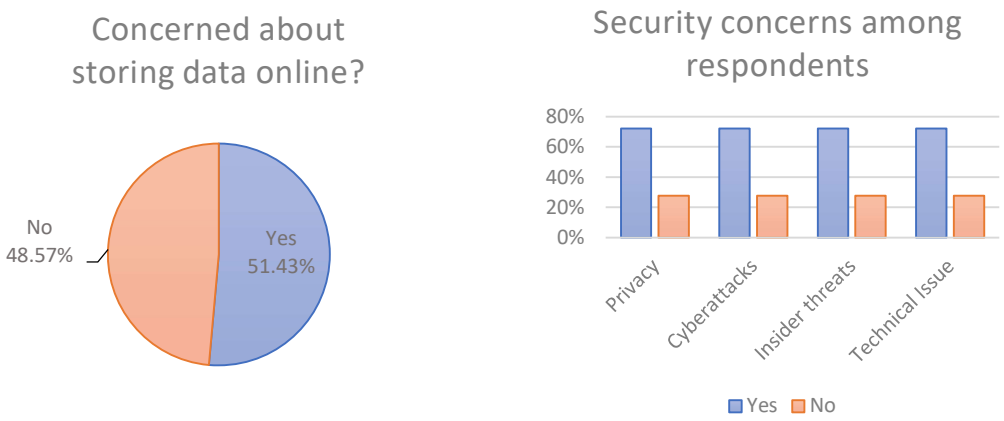

Figure 7. Respondents' concerns about storing data online

As shown in Figure 7, more than half of respondents have concerns about using and storing data on a cloud-based software. However, their concern is spread equally among the four types of threats though the risk of happening can vary, for example, in practice and considering all data protection regulations, technical issues are more probable than cyberattacks or insider threats. It can be argued that the respondents have no specific concern about using a cloud-based solution, but more likely they are hesitant, resistant against the concept of cloud itself for what they have been hearing about it. This argument is partially confirmed through the next survey question.

e. Concerns about storing research data online do not prevent most users from adopting a cloud-based CAQDAS.

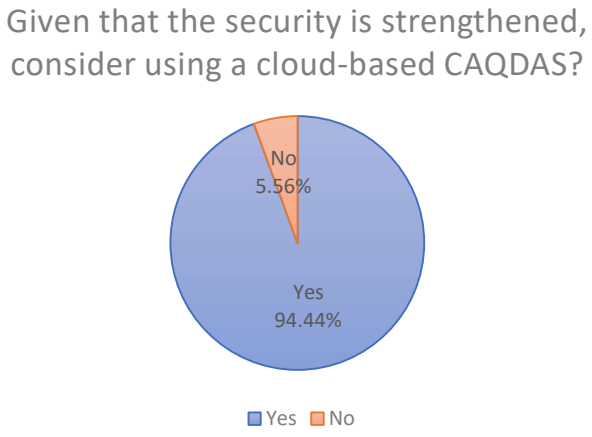

Figure 8. Respondents' willingness to adopt a cloud-based CAQDAS

As indicated in Figure 8, around 94 percent of respondents with security concerns mentioned that they would still consider using a cloud-based CAQDAS if the security is ensured and strengthened. "Ensured" and "strengthened" security is only a relative

concept because something can be more secured to some people but not to all others. Thus, it strengthens the argument mentioned earlier that the respondents may have vague concerns about security of cloud computing and cloud-based solutions. They can very likely be swayed to adopt a cloud-based CAQDAS if it is clear to them that their concerns are soothed, through additional security measures like backup, restoration, autosave, link or integration to personal drive, secure payment protocols and tighter access control (e.g. two-step verification).

f. Potential users who are willing to pay prefer a rolling subscription.

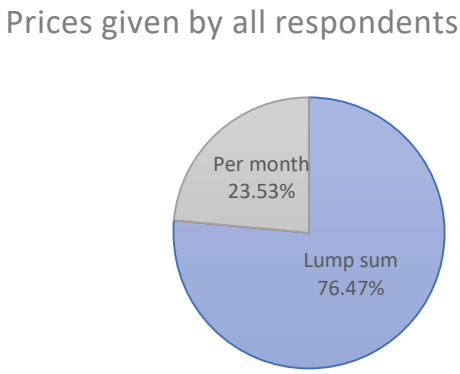

Figure 9. How respondents quantify their willingness to pay

According to Figure 9, respondents tend to provide one-time prices (around 76 percent of all), probably because it is easier to for them to tell and quantify the maximum price they are willing to pay in one lump sum, while recurring fees are more difficult to control and estimate over time in terms of total incurred costs.

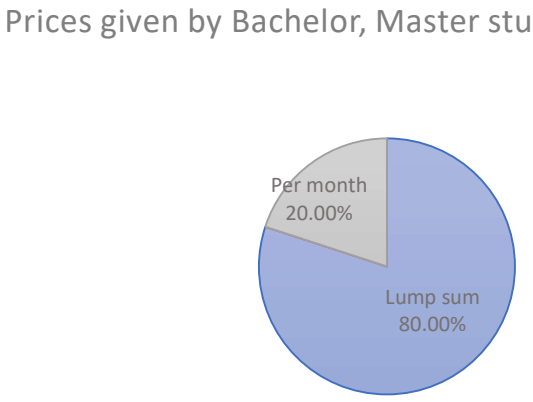

Figure 10. Different forms in which Bachelor, Master students give prices

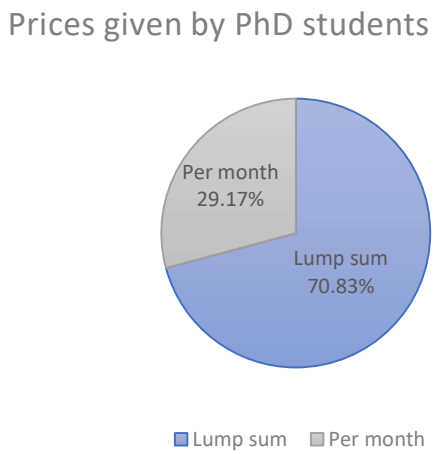

Figure 11. Different forms in which PhD students give their prices

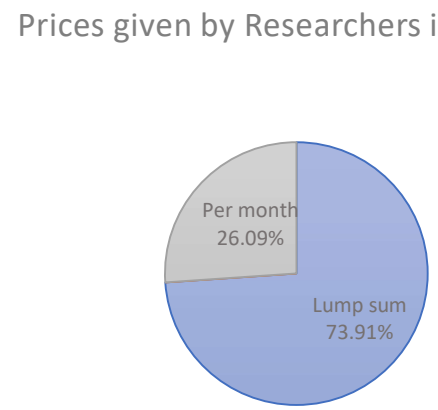

Figure 12. Different forms in which Researchers give their prices

Inferred from Figure 10, 11, 12, there is not much a difference among the three groups of users. Specifically, 80 percent of Bachelor, Master students, around 71 percent of PhD students, and 74 percent of researchers provided their prices as a one-time fee.

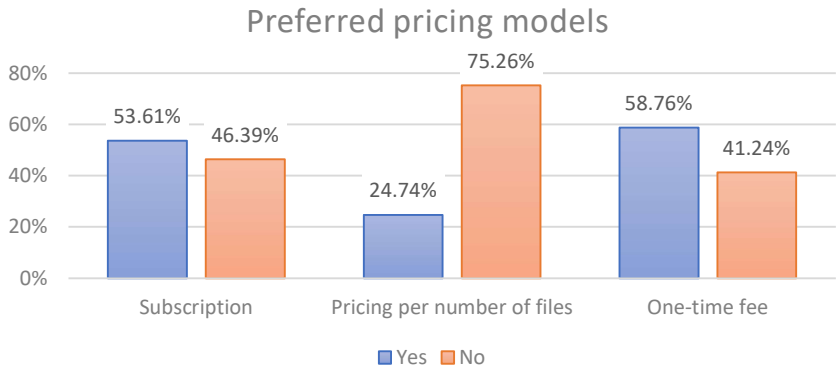

Figure 13. Preferred Pricing models across all user groups

When asked which pricing model makes sense, there is not a big difference between Subscription (53.61% yes) and One-time fee (58.8% yes), as shown in Figure 13. Pricing per number of files (24.74% yes) appears not that appealing to respondents.

One-time fee refers to a perpetual license, however, even in this case, the fee is never paid once throughout the lifetime of the service. There are also upgrade fees once every several years and this, upon revelation to users, may hinder them from paying upfront a large lump sum while having to pay later to keep the software up to date. Subscription refers to recurring payments as the service continues. This pricing is charged every month, every 6 months, a year or two years. More than half of respondents opt for this pricing model, meaning it is a viable option for final pricing of QDAcity.

All in all, potential users tend to be indifferent between subscription and one-time fee. However, they show clearly their preference for a more predictable type of cost, through their choice of pricing models and how they quantify their willing-to-pay prices.

3) Functionality preferences:

g. Offline use is an important feature for a CAQDAS.

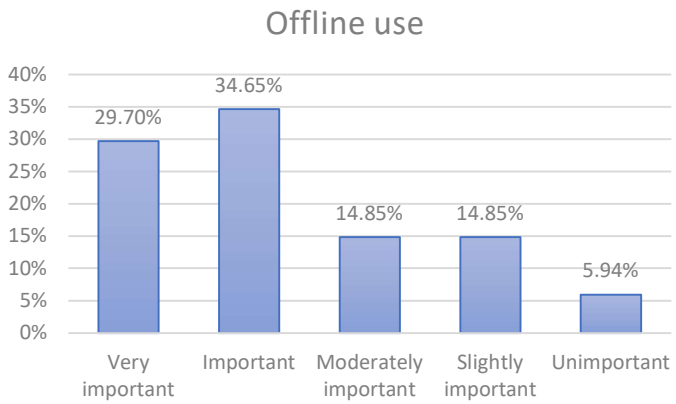

Figure 14. Rating of "Offline use" feature

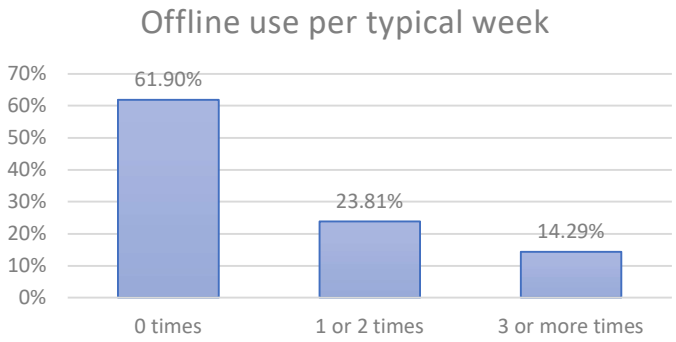

Figure 15. Offline use among "Important", "Very Important" raters

From Figure 14, it is clear that around 64 percent of respondents rate this feature as either “Important” or “Very Important”. However, among these, more than 61 percent report no offline use in a typical week (Figure 15). Offline use is still a reasonable feature, for more than 38 percent of respondents who need to work offline from at least once to thrice per week. However, considering all data options people can get easily nowadays and they may need Internet also for other purposes, offline use may not be that critical to them and thus not for QDAcity.

h. Real-time group coding is an important feature for a CAQDAS.

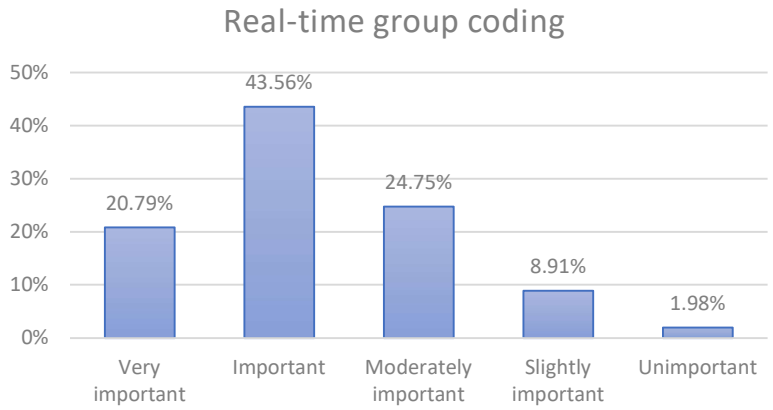

Figure 16. Rating of “Real-time group coding” feature

According to Figure 16, around 43.56 percent of the respondents think this feature as “Important” and 20.79 percent as “Very important” (a total of around 64.35 percent).

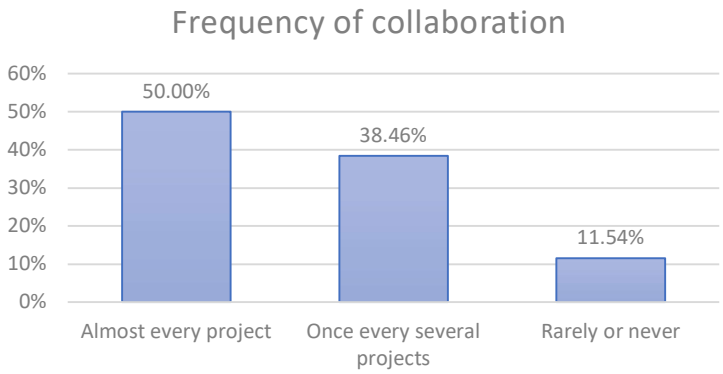

Figure 17. Frequency of collaboration among “Important”, “Very Important” raters

According to Figure 17, among respondents who rate this feature as “Important” or “Very important”, around 88.46 percent claimed to collaborate with others once in every several projects or in almost every project.

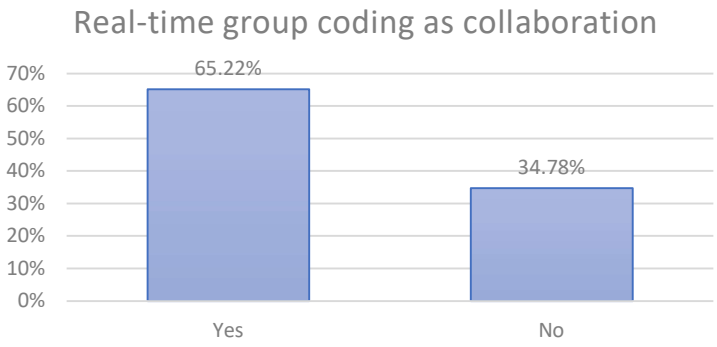

Figure 18. Real-time group coding in collaboration

Figure 18 shows a more drilled-down result from Figure 17. More than 65 percent of respondents who collaborate in almost every project or once every several projects actually do real-time coding as part of their collaboration. This would be equal to around 57 percent of all responses who rate this feature as important and actually use it. Since the figure is not dominating, rather than critical, this feature is only nice to have for a CAQDAS.

i. Measure of interrater reliability is an important feature for a CAQDAS.

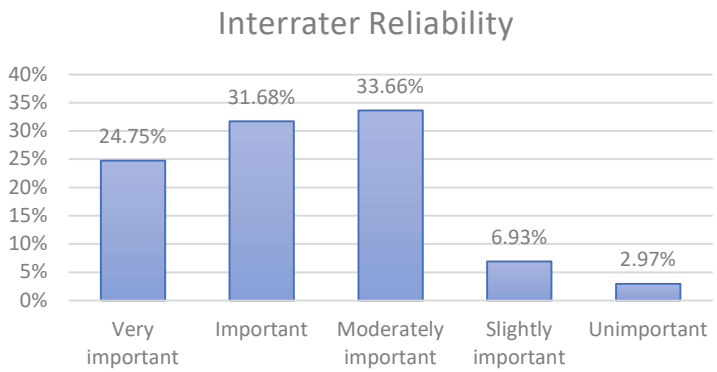

Figure 19. Rating of "interrater reliability" feature

According to Figure 19, only around 56 percent of respondents rate this feature as "Important" or "Very Important".

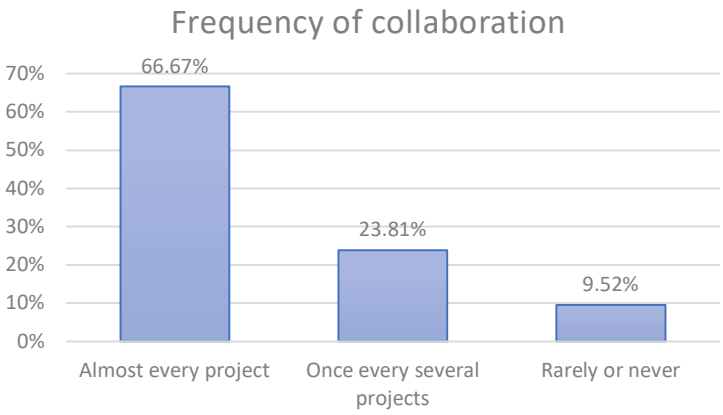

Figure 20. Frequency of collaboration among "Important", "Very Important" raters

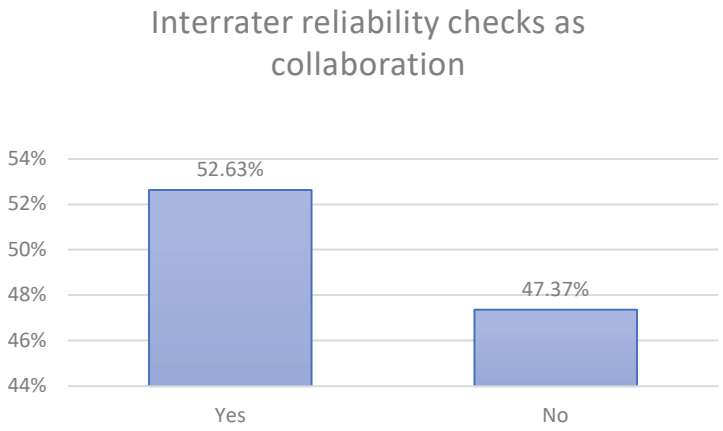

Figure 21. Interrater reliability in collaboration

As shown in Figure 20 and 21, out of the respondents who rate this feature as "Important" or "Very important", 66.67 percent collaborate with others in almost every project and 23.81 percent at least once every several projects, a total of up to 90 percent. Among those, 52.63 percent, use interrater agreement in collaboration. However, since only a little more than half of these potential users actually use the feature, it is not critical for a CAQDAS, but nice to have.

j. Support for member-checking (permission schemes) is an important feature for a CAQDAS.

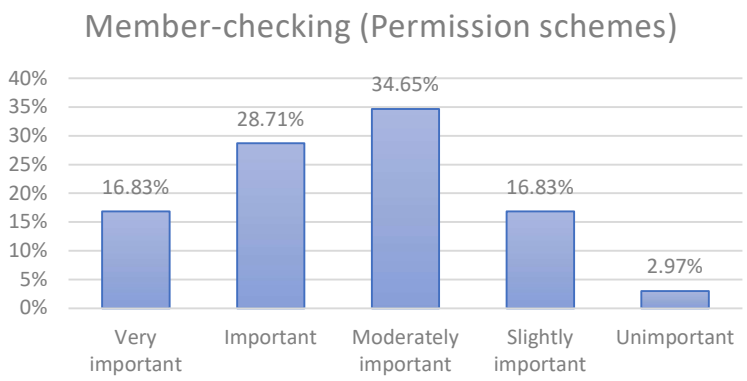

Figure 22. Rating of "Member-checking (permission schemes)" feature

As shown in Figure 22, more than 54 percent of respondents consider this feature "Unimportant", "Slightly important" or "Moderately important".

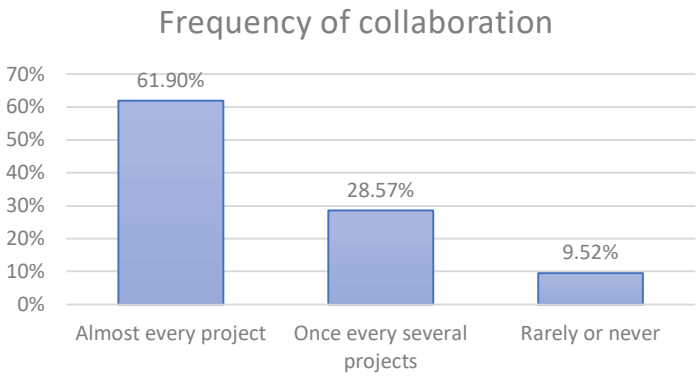

Figure 23. Frequency of collaboration among "Important", "Very Important" raters

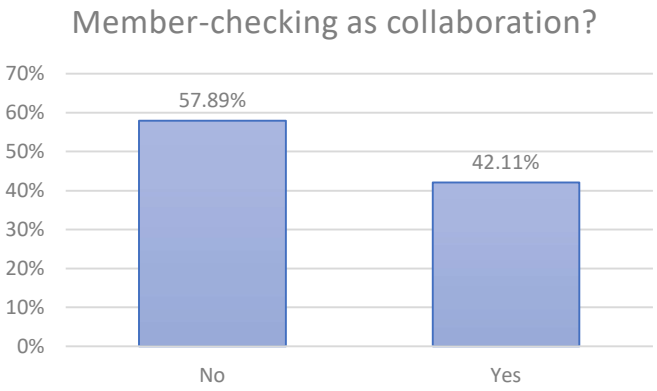

Figure 24. Member-checking in collaboration

Among approximately 46 percent of respondents who consider this feature "Important" or "Very important", 61.90 percent collaborate in almost every project, 28.57 percent once every several projects but only 42.11 percent out of all these do member-checking as part of their collaboration. This feature is not needed that frequently and therefore, is only nice to have but not critical for a CAQDAS.

k. Transcribing audio/video files is an important feature for a CAQDAS.

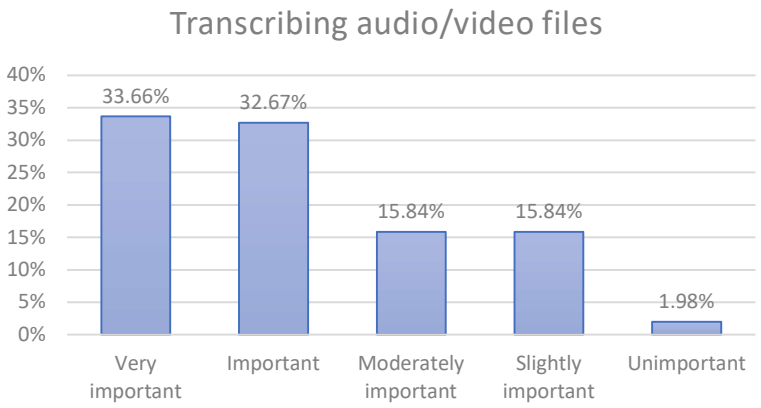

Figure 25. Rating of "Transcribing audio/video files" feature

According to Figure 25, more than 66 percent of respondents think this is "Important" or "Very important".

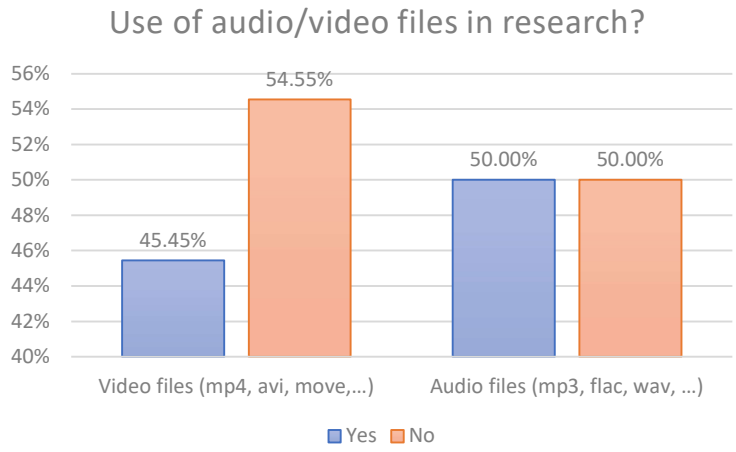

Figure 26. Use of audio/video files among "Important", "Very important" raters

Among this 66 percent, as shown in Figure 26, 50 percent actually uses audio files in their research. Meanwhile, only up to 45 percent uses video files. Accordingly, transcribing of audio files is potentially more of demand than transcribing of video files. However, without further data, the numbers are not that significant to confirm any solid demand for transcribing feature in QDAcity. This feature, therefore, is a great feature to include but not critical for a CAQDAS. If it were to be included in QDAcity, transcribing of audio files should be prioritized to that of videos.

l. Coding of audio/video files is an important feature for a CAQDAS.

Coding of audio/video files is different from transcribing. Coding of audio or videos involves direct categorizing and thematic labeling of different parts across the length of the media file. This feature is not a standard one in a CAQDAS but an advanced one that has been available for some time in the market, e.g. in NVivo. The technology for this feature thus is supposed to be more sophisticated than transcribing.

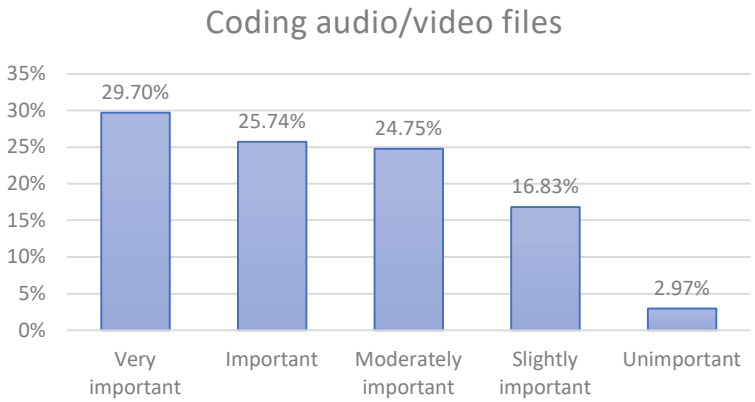

Figure 27. Rating of "Coding audio/video files" feature

As shown in Figure 27, more than 55 percent of the respondents think this feature "Important" to "Very important".

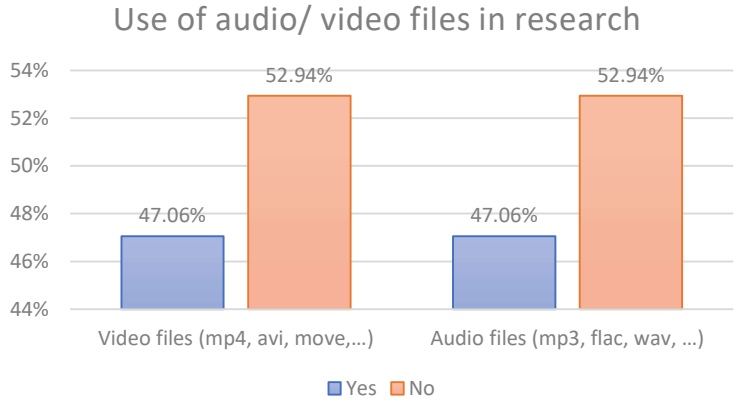

Figure 28. Use of audio/video files among "Important", "Very important" raters

Drilling it down as in Figure 28, among these respondents, only 47 percent actually uses video files or audio files in their research activities. This is therefore not a critical but nice to have feature. Considering the required effort to implement such a feature and little demand observed for use of media files, it may not be practical to include. In addition, with transcribing feature and generated transcripts, text coding is always possible and analysis is good to go without direct coding of multimedia files. Thus, this feature can be considered unnecessary for a CAQDAS.

m. Support for processing multiple file types is an important feature for a CAQDAS.

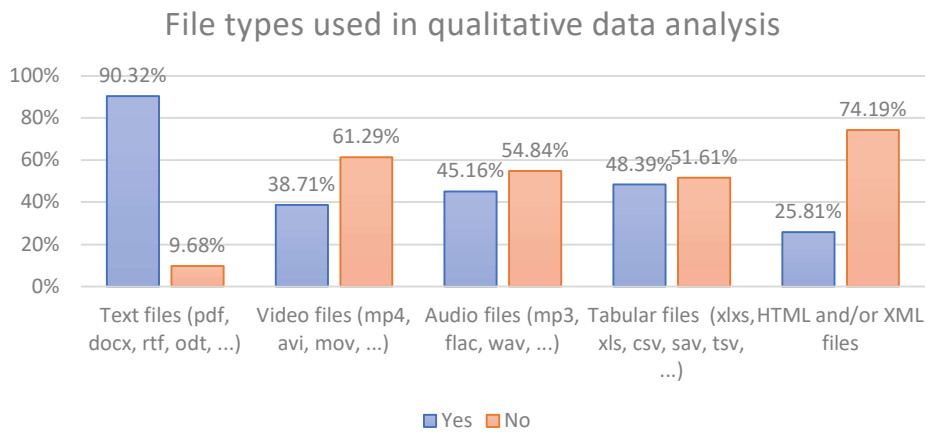

Figure 29. File types used in qualitative data analysis by all respondents

Among all responses, more than 90 percent use text files (pdf, docx, odt, rtf, txt, …) in their qualitative research and data analysis, as shown in Figure 29. Thus, there is indisputable demand for use of text files in a CAQDAS.

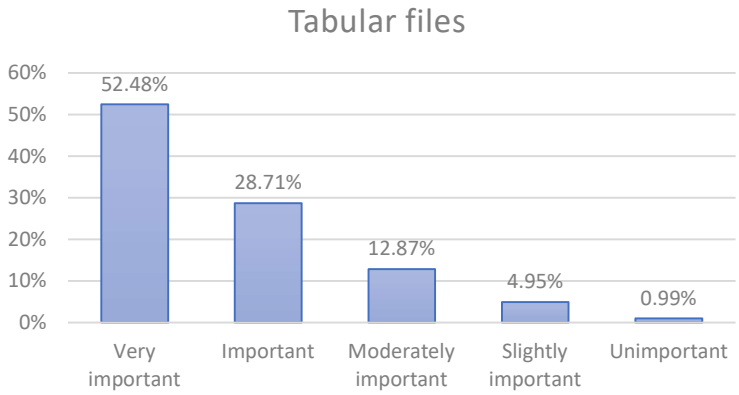

Figure 30. Rating of "Tabular files" support

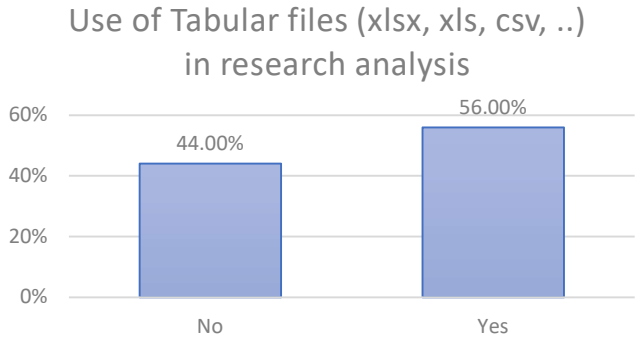

Figure 31. Use of Tabular files among "Important", "Very important" raters

More than 80 percent of all respondents who answered this question think that support for tabular files is “Important” to “Very important”, as indicated in Figure 30. Drilling down this 80 percent, 56 percent actually use tabular files in their research analysis, as shown in Figure 31. This indicates potential demand for processing of tabular files in CAQDAS. Considering also the popularity of excel, csv file types in textual data handling and office work, support for these files should be critical for a CAQDAS.

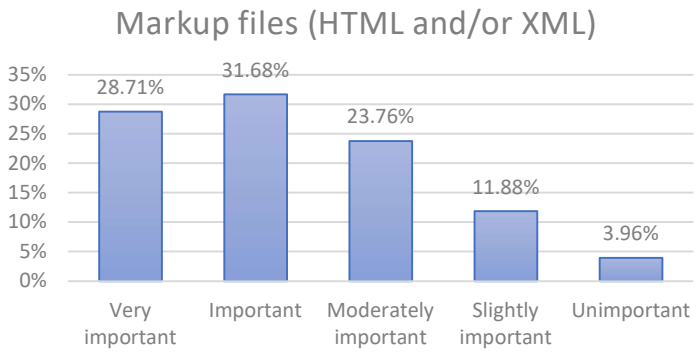

Figure 32. Rating of “HTML and/or XML files” support

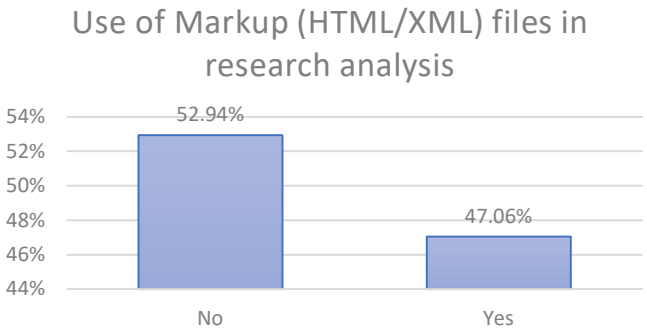

Figure 33. Use of Markup files among “Important”, “Very important” raters

As demonstrated in Figure 32, more than 60 percent of respondents consider processing of markup files (HTML, XML) “Important” to “Very important”. Within these groups of respondents, as illustrated in Figure 33, only up to 47 percent – less than half - actually use such files in research activities. This means no solid demand observed for support of these file types in a CAQDAS and therefore, may not be beneficial to include in a CAQDAS.

n. Saturation measure is an important feature for a CAQDAS

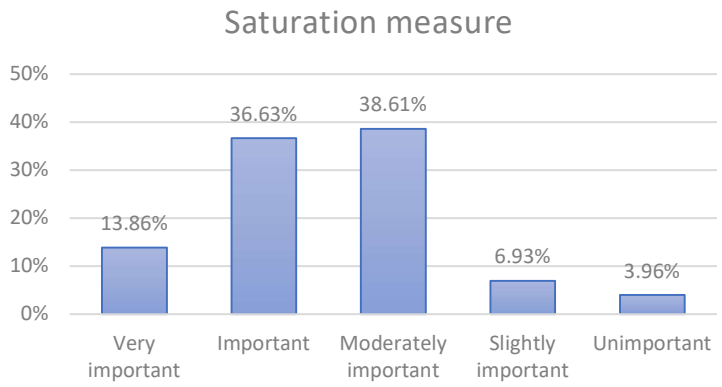

Figure 34. Rating of "Saturation measure" feature

According to Figure 34, more than half of the responses claim saturation measure is "Important" or "Very important". However, since the difference in proportions is very slight, saturation measure can be considered only a nice-to-have.

o. Professional technical support (e.g. call center) is an important feature for a CAQDAS.

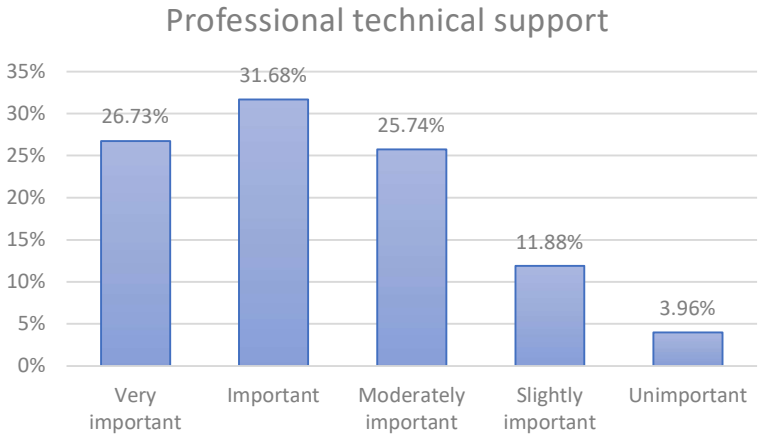

Figure 35. Rating of "Professional technical support" feature

According to Figure 35, a total of around 58 percent of all survey respondents thinks this feature is either "Important" or "Very important". Since up to 42 percent of responses state professional technical support is less than important, this service is not critical for a CAQDAS but can be nice to have. Professional support here does not necessarily involve offering 24/7 service at a very high cost. Rather, it can be offered in the form of call service during working hours.

p. Dictionary is an important feature for a CAQDAS.

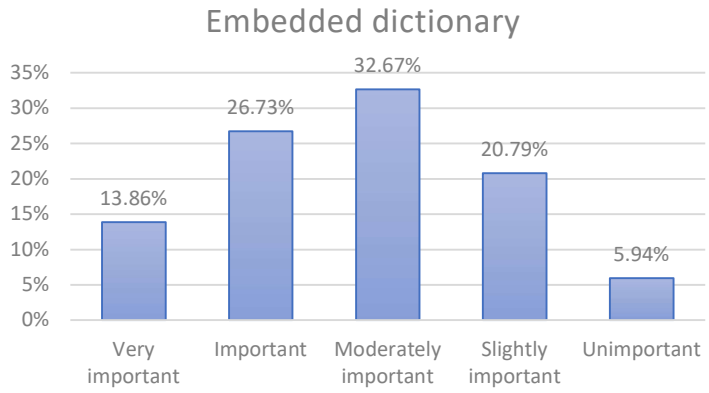

Figure 36. Rating of "Dictionary" feature

Figure 36 shows that around 60 percent of respondents consider this feature less than important. It is, as a result, not considered beneficial for a CAQDAS within the scope of this thesis.

q. Mobile version or portability is an important feature for a CAQDAS.

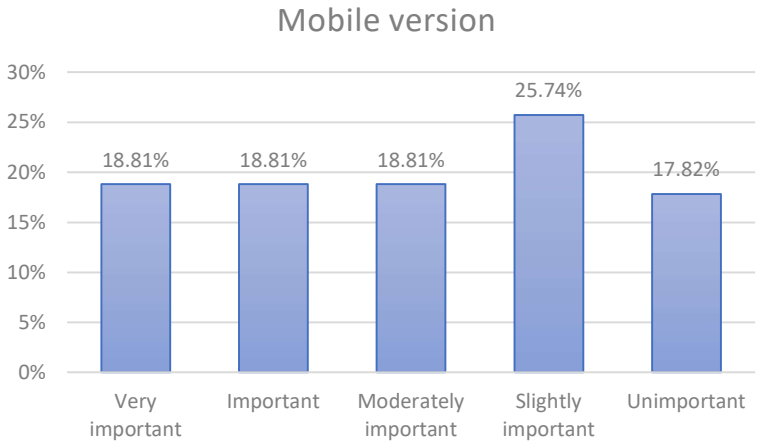

Figure 37. Rating of "mobile version" feature

As shown in Figure 37, more than 62 percent of respondents thinks this feature is less than important. Without further evidence, this feature is not necessary for a CAQDAS and thus will not be considered for a CAQDAS in the scope of this paper.

r. Autosave, versioning and backup options are important features for a cloud-based CAQDAS

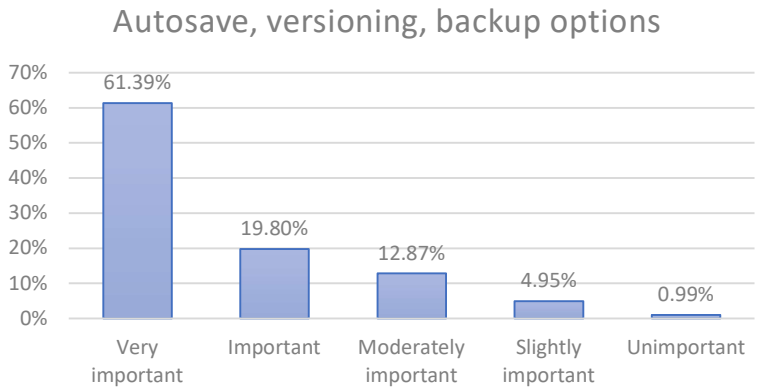

Figure 38. Rating of "Autosave, versioning and backup options"

Among all responses, as indicated in Figure 38, more than 61 percent rate this feature group as "Very important" and an additional 20 percent as "Important". This also aligns with the potential users' concerns of security. Within this 81 percent of responses, around 80 percent expressed concerns about having technical issues with a cloud-based software and 70 percent about cyberattacks, both causing interruptions in services and loss of data.

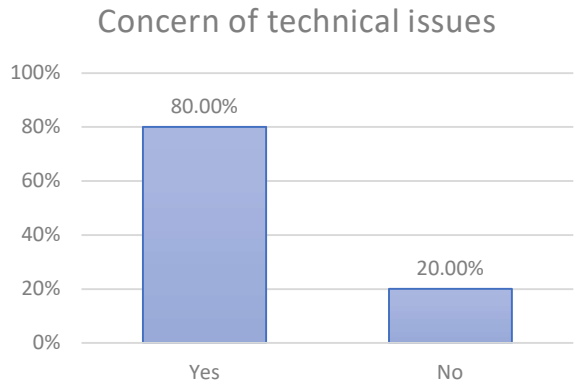

Figure 39. Concern of technical issues among "Important", "Very important" raters

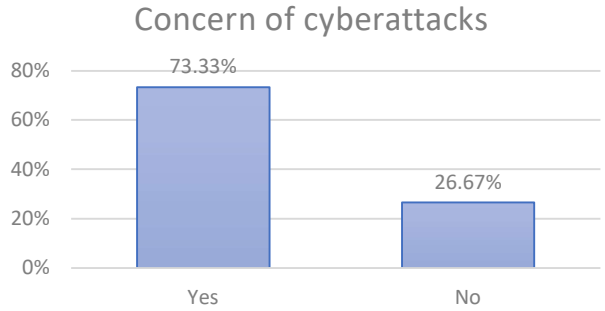

Figure 40. Cyberattack concern among "Important", "Very important" raters

Due to dominating figures for “Yes” in Figure 39 and 40, backup and versioning are of no doubt critical features for a cloud-based CAQDAS.

s. There are some other important features for a cloud-based CAQDAS.

Table 3 is the list of valid answers that survey respondents provided for the question “Do you think of any other feature that is important for a CAQDAS?”. These answers are retrieved from a total of 141 survey responses including those that are not complete, since the incomplete ones can also serve to provide useful insights. In the second column, there is grouping information for these answers which allows determination of proportions of respondents with similar ideas.

Answers	**Feature group**
Option for multiple licenses with discount (for research teams).	Not a feature
Create codes to identify patterns	Automatic code identification
Instructions for data analysis step by step	Tutorial, training
being able to create graphs out of the data	Visualization
A one-month free test version.	Not a feature
Should be easy to use	Not a feature
Really good docs, in case this is not meant by the embedded dictionary question.	Not a feature
Mindmapping	Visualization
Collaboration of multiple users	Collaboration
Absolute Reproducibility	Interrater Reliability
Ease of use	Not a feature
The coding itself has to be ON POINT.	Not a feature
Speed	Not a feature
Some sort of feedback mechanism from the user would be good.	Support for Member-checking
Some kind of diagrams or overview of e.g. coded sections	Visualization
Visualization Capabilities. Needs to be "fun".	Visualization
Not exactly a feature, but more like how easy it is for new users? Is it easy to learn how to use the CAQDAS? Is the UI/UX self-explanatory? If there is a missing feature, how fast the development team could support me to implement it / guide me to customize the platform to fit my need?	Not a feature
It is really basic but a really simple user interface is always important.	Not a feature
Connection to database systems	Integrations
Quickly access	Not a feature
Automatically suggesting commands or statements	Automatic code identification
Detailed user instructions	Tutorial, training
Example projects, best practice guidelines	Tutorial, training

An interface for import data manually would be nice	Manual import
Facile ways of maneuvering “saturation” to be part of the process - rather than drawing it out on paper.	Saturation
Not just versioning, but control over who can access what. Memoing along the way, ability to notate changes, so documentation of the project during progress	Memo, documentation
Option to create word clouds or visual representation of the data that you coded	Visualization
Different languages!!!	Multi-lingual
Simple interface, easy to navigate, DOI-Import	Import work using DOI
Speech recognition	Speech recognition
Graphical options, e.g. plotting codes on a graph or as bigger and smaller circles based on the amount of mentiones in a file or dataset.	Visualization
Ai learning	Automatic code identification
User interface design -- the more clear the UI the better.	Not a feature
Visualization	Visualization
Tools to support the development of models, e.g., frequency analysis, visualization	Visualization
EU-local data storage / processing	Not a feature
I am not familiar with this but if I have to write something maybe a feature to save/export some data in multiple formats	Multi-format support
it should be user friendly and therefore intuitive	Not a feature
User friendliness.	Not a feature
Share results with other or to a public audience	Support for Member-checking
Visual representation of the coding scheme.	Visualization
Not crashing	Not a feature

Table 3. Additional features respondents deem important for a CAQDAS

Among these 42 answers, only 27 address a specific (or two) feature. Among these 27 ideas of features, 9 are about visualization options, 3 about automatic pattern and code identification, 3 about tutorials or training, and some individual mentions of interrater reliability, memo-ing, member-checking support, multi-lingual, saturation and multi-format support. Besides those that overlap with features already asked in survey questions, visualization and automatic code identification are the two new ones. With 9 mentions in total out of 27 (more than 30 percent), written down specifically by respondents, visualization is an important feature to consider for a CAQDAS. Tutorial or training is also a critical part of a CAQDAS or any tool in general as it ensures users are well informed of how to navigate through and work with the software. 3 out of 27 respondents also list “Tutorial, training” as an important feature for a CAQDAS, as shown in Table 3.

As for the other ideas like memo-ing, group pricing (for multiple users), multilingual

support or import documents using DOI, the development team can consider implementing them later if further validation confirms they are actually in demand.

In addition, three survey respondents filled in the "Other" option of question "Which file types do you work with?" with json, images, .py and .cpp. While .py and .pp are programming files and may not be relevant for qualitative data analysis. "images" is a legit suggestion. Though "images" is a bit too general, the most popular image file extensions like jpeg., png., bmp., … can still be considered for implementation. JSON files are also widely used in qualitative research, e.g. exports from survey tool can be in JSON format. Thus, it is also a valid suggestion.

However, as these file types are mentioned once only among all survey respondents, it is hard to confirm any criticality. The development team can consider implementing JSON and image file formats later if the demand for these is proved.

t. There is potential demand for integrations between CAQDAS and some third-party applications.

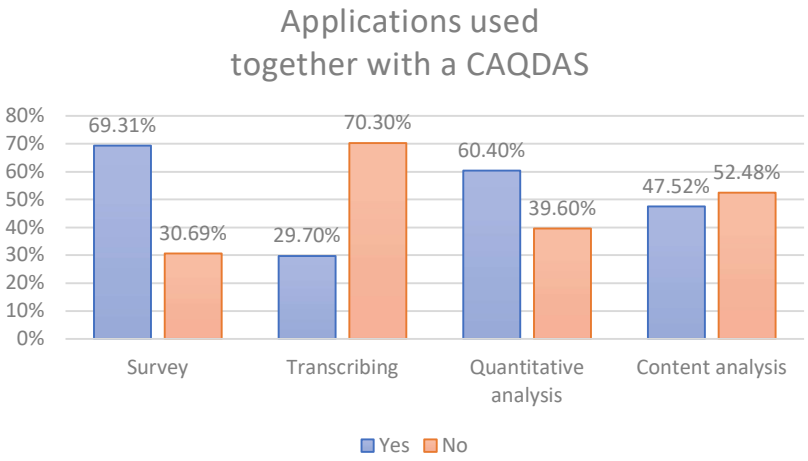

Figure 41. Types of applications used together with CAQDAS

From Figure 41, it can be seen that there is potential demand for use of CAQDAS together with:

1) Survey tools
2) Quantitative data analysis tools

This promises also demand for integrations between these solutions. Therefore, it is beneficial to have such integrations in a CAQDAS.

5.1.3 Summary of hypotheses

Table 4 summarizes the conclusions to the hypotheses regarding market potential, product proposition, pricing and preferred features for a CAQDAS. This information will be used to craft the marketing plan for QDAcity in the next chapter.

Hypothesis	Conclusion based on survey results
There are potential users who are willing to pay for personal use of CAQDAS.	True, 34.65 percent of respondents are willing to pay for personal use of CAQDAS.
There is demand for a lightweight and focused CAQDAS.	True, 29.41 percent of respondents prefers a lightweight and focused CAQDAS.
A majority of students would opt for a free option if they used a CAQDAS.	A majority of students would be willing to pay at least something for a CAQDAS.
There are potential users who have concerns about storing their data online.	True, there is also an equal distribution among all four types of concerns: privacy, cyberattacks, insider threats, and technical issues.
Concerns about storing research data online do not prevent most users from adopting a cloud-based CAQDAS.	True, 94 percent of those who have concerns are willing to use a cloud-based CAQDAS given better security.
Potential users who are willing to pay prefer a rolling subscription.	There is almost equivalent preference between subscription and one-time fee. Respondents tend to quantify their willingness to pay in a lump sum.
Offline use is an important feature for a CAQDAS.	Offline use is not necessary for a CAQDAS.
Real-time group coding is an important feature for a CAQDAS.	Real-time group coding is nice to have for a CAQDAS.
Measure of interrater reliability is an important feature for a CAQDAS.	Measure of interrater reliability is nice to have for a CAQDAS.
Support for member-checking (permission schemes) is an important feature for a CAQDAS.	Support for member-checking (permission schemes) is nice to have for a CAQDAS.
Transcribing audio/video files is an important feature for a CAQDAS.	Transcribing audio files is nice to have for a CAQDAS. Transcribing video files is however not necessary.
Coding of audio/video files is an important feature for a CAQDAS.	Coding of audio/video files is not necessary for a CAQDAS.
Support for processing multiple file types is an important feature for a CAQDAS.	Support for text files and tabular files are important for a CAQDAS.

Saturation measure is an important feature for a CAQDAS.	Saturation measure is nice to have for a CAQDAS.
Professional technical support (e.g. call center) is an important feature for a CAQDAS.	Professional technical support (e.g. call center) is nice to have for a CAQDAS.
Dictionary is an important feature for a CAQDAS.	Dictionary is unnecessary for a CAQDAS.
Mobile version or portability is an important feature for a CAQDAS.	Mobile version or portability is unnecessary for a CAQDAS.
Autosave, versioning and backup options are important features for a cloud-based CAQDAS.	True, autosave, versioning and backup options are important for a cloud-based CAQDAS.
There is demand for integrations between the CAQDAS and some third-party applications.	There is observed demand for use of a CAQDAS together with survey tools and quantitative data analysis tools. This therefore promises also demand in integrations between a CAQDAS and these types of software.
Additional preferred feature(s) or pricing	Visualization options and tutorials are critical features for a CAQDAS. Automatic code recognition, memo-ing, multilingual support, group pricing, import documents using DOI, support for JSON and image files can be promising for future implementation.

Table 4. Summary of conclusions to hypotheses

5.2 Market competition

5.2.1 Competitor Analysis

This section is to present the leading products in the market of CAQDAS and what features, pricing they offer. Such details, as shown in Table 5, could serve as a benchmark for determining what to offer and what not in QDAcity.

Competitor / **Criteria**	**MAXQDA** (*MAXQDA*, n.d.)	**Atlas.ti** (*ATLAS.ti*, n.d.)	**Nvivo** (*NVivo*, n.d.)	**Taguette** (*Taguette*, n.d.)	**Dedoose** (*Dedoose Features*, n.d.)	**QDAminer 6** (*QDAMiner Features*, n.d.)	**QDAminer Lite** (*QDAMiner Lite Version*, n.d.)
Cloud-based?	No	Cloud (web) version with limitations of features	Transcribing and Collaboration	No	Yes	No	No

		Link	are Cloud features.				
Free trial?	Yes (30 days)	Yes (5 days – Cloud)	Yes (14 days)	N/A	Yes (30 days)	Yes (30 days)	N/A
Key features	Coding Categorizing Text analysis Member-checking or sharing of results (MaxQDA Reader) Visualization: Code distribution matrix, overlapping codes, word cloud, mind mapping, … Mixed-method analysis Dictionary Memos, annotations Multi-lingual coding Transcribing Statistical analysis (Pro Analytics)	Coding Categorizing Auto-coding Structure recognition Analysis and visualization: network map, hierarchical clustering, bar graph (frequency), … Real-time collaboration User management Querying, memos, annotations Within and cross case analysis Interrater reliability Sentiment analysis Multi-lingual UI	Coding Categorizing Analysis and visualization: code distribution matrix, modeling, frequency, clustering, word tree, tree map, word cloud, … (Jew, 2014) Interrater reliability Automatic identification of themes Suggestion of insights and queries Group use (for extra fee) Transcribing (for extra fee) Coding of audio/videos (Pro and Plus version)	Coding Categorizing Group real-time coding (web) User management	Coding Categorizing Real-time coding User management Analysis and visualization: bar graph (frequency), co-occurrence matrix, … Multiple analysis methods, including mixed-method Interrater reliability Coding of video/audio Multi-lingual coding Memo system Trainings/ professional support	Coding Categorizing Analysis and visualization: tag clouds, 2D map, correspondence analysis plot, … Cluster Extraction, drag-and-drop coding Code Similarity Query by Examples Geographic and time analysis Statistical analysis Multi-user settings Interrater reliability Multi-lingual UI	Coding Categorizing Overview graphs, charts of coded sections / keywords Multi-lingual UI
Supported file types	Imports: txt, PDF, survey, video and graphical files Exports: docx, xlsx, PDF, images (png, svg), HTML, RTF, .MEX (coded segments), txt and tab-delimited	Imports: doc, docx, txt, (exc. tables and images, graphical), PDF Exports: codebook, xls, xlsx., doc, docx, PDF, XML, QDPX (project), to statistical software	Imports: survey files, from any source and most types Exports: doc, docx, xls, xlsx, PDF, RTF, txt, HTM, HTML, jpg, audio/video files (in original formats)	Imports: IPDFs, Word Docs (.doc, docx), Text files (.txt), HTML, EPUB, MOBI, Open Documents (.odt), and Rich Text Files (.rtf) Exports: highlighted texts	Imports: doc, docx, txt, PDF, HTML, images: jpgs, gifs, etc., audio/video, streaming files, xls, xls, xlsx, csv. Exports: doc, docx, xls, xlsx, codebook, QDPX (project) Migrations from/to other CAQDAS are supported.	Imports: plain text, RTF, HTML, PDF, xls, xlsx, csv, tsv, Access, from statistical tools (SPSS, Stata), social media (Facebook, Reddit, Youtube, …), emails, surveys (Qualtrics, SurveyMonkey,…), graphics (bmp, jpg, png,…), XML, reference information system (.ris) files Exports: doc, docx, xls, xlsx, codebook, images (for charts, BMF, png,…), CSV, HTML, XML, and delimited ASCII files.	Imports: plain text, RTF, HTML, PDF, xls, xlsx, csv, tsv, Access, tab-delimited, imports from other CAQDAS, transcription tools, reference information system (.ris) files Exports: xls, xlsx, xls, Tab Delimited, CSV, doc, docx, bmp, png, jpeg, wmf, QDP (project)

Integration	N/A	N/A	Citavi (citation service)	N/A	N/A	SimStats (Statistical analysis) WordStats (Text analysis and mining)	N/A
Pricing - Students	- MaxQDA Reader: free - Standard: N/A - Plus: $47/6 months $95/2 years - Pro & Analytics: $55/6 months $110/2 years	- Web, PC, Mac: $99/2 years $51/6 months $10/month	$81/year (limited for 12 months only) ~ €5.83 / month	Free	$10.95/month	$595 (perpetual) $238/year $295 (upgrade)	Free
Pricing – Non-students	- MaxQDA Reader: free - Standard: $380 (annual) $1,140 (perpetual) $570 (upgrade) - Plus: $460 (annual) $1,380 (perpetual) $690 (upgrade) - Pro & Analytics: $499 (annual) $1,499 (perpetual) $749,50 (upgrade)	- Only Web: $50/month (commercial) $20/month (educational) - Web, PC, Mac: + Individual: $600/year $1,840 (perpetual) + Group use: 5 users: $3,100 (annual) $6,800 (perpetual) 10 users: $5800 (annual) $13,200 (perpetual)	- Academic: + Windows: €696 ($807) + Mac: €553 ($641) - Non-academic: + Windows: €1024 ($1188) + Mac: €784 ($909) All are perpetual licenses	Free	- Individual: $14.95 - Group use: + 6+ users: $10.95/user/ month + 2-5 users: $12.95/user /month	$2,595 (perpetual) $1,038/year $1,295 (upgrade)	Free

Table 5. Summary of key features and offers of leading competitors

5.2.2 Unique selling points

Based on the competitor analysis in Table 5 and existing or planned features of QDAcity, unique selling points of QDAcity can be inferred as following.

Among the popular competitors as mentioned above, most are feature-rich with a wide range of functionalities and options for data processing, analyzing, collaborating, visualizing, as well as importing, exporting in multiple formats. Two out of these, on the other hand, are extremely light software with no cost at all. People might like a lot of features all at once place, as shown in their responses to the question about "Type of CAQDAS"; however, they can also easily get confused when using it or find it unaffordable, unless they have a very sophisticated, extensive need in working with qualitative data. Meanwhile, pure coding software is easy to use, free

of charge, can be used for a wider range of purposes but then may not cover some emerging needs of processing a specific file type, collaborating measures, or visualizations. Maybe due to this reason, only 29 percent of the respondents claim to prefer a lightweight and focused solution. This hints a niche opportunity for a solution in-between, which QDAcity can take advantage of.

In addition, those features of QDAcity that are not offered by the competitors can also serve as its unique selling points. These include saturation rate and visualization of interrater agreement (agreement map).

Equipping QDAcity with additional features should be done moderately; otherwise it would also get closer to becoming a feature-rich solution. Accordingly, the number of added features should be limited, and only those that present to be clearly important, needed and helpful for a majority of potential users would be considered.

6 Discussion and Implications

The survey results and earlier derived implications on the importance of functionalities as well as pricing serve the base for specific business recommendations for QDAcity. These recommendations are consolidated into a marketing strategy, based on which sales and revenue estimations for QDAcity will also be calculated.

6.1 4P Strategy

4P strategy is a marketing strategy that defines the proposition, pricing, placement (or distribution), and promotion of a product, in this case QDAcity.

- **Product**

In the survey, respondents are asked to choose their preferred type of CAQDAS. Among the responses, only 29 percent choose a focused, lightweight solution while more than 70 percent vote for a feature-rich one. It may not therefore be beneficial for QDAcity to stay completely with the "focused, lightweight" concept as the development team planned initially. However, QDAcity should not lean more towards a "feature-rich" or end-to-end concept either as there are already solutions out there which have been offering the same for some time and leading the market. As shown in the section Market Competition, an "in-between" solution is not fully exploited in the CAQDAS market. Thus, it can be an opportunity, a direction for QDAcity.

Users' needs do not stay intact over time. For example, a student who needs a simple coding tool for his thesis may end up needing more extensive functionalities for his research after he graduates and starts working as a researcher. Thus, QDAcity may better offer at least two tiers of functionalities:

- Free version
- Proprietary add-up for some additional features

Some students and even researchers, depending on their purposes, may choose the free version for simple thematic coding and categorizing. If they need additional features, they can opt to pay for those. The free version offers users the chance to try out some basic functionalities at no cost, to store their data in QDAcity. It can be considered a vendor lock-in when users first choose QDAcity for its free version but then find it cumbersome to move their data to another solution. In case of more needs, they may be more tempted to choose to upgrade their account with QDAcity. However, to keep the software expansion from a feature creep, there should not be a whole lot of features to be added. Only those that are uniquely available in QDAcity, considered important by potential users and facilitate easier coding or analysis process rather than end-to-end data handling should be selected. The cost for these additional features, therefore, is supposed to be significantly lower than the costs they have to pay for a feature-rich software in the market.

A more detailed offer of features for the two versions is recommended in Table 6,

considering the existing functionalities of QDAcity, existing offers on the market, and survey results (i.e. potential users' opinions). After all, it makes sense to include in the free version the critical features of a CAQDAS which QDAcity has yet to offer because they are already offered in other free solutions in the market. Otherwise, users may simply choose Taguette or QDAminer Lite in the beginning for a more complete package at no cost. Proprietary add-on will include those features concluded in the earlier chapter as nice-to-haves.

Though the earlier results show that users tend to use CAQDAS with survey and quantitative analysis tools, it is not clear in specific which tools. If the development team decides to pursue an integration of QDAcity with a survey or quantitative analysis tool (or both), it can be helpful to determine on the most popular names or those that have the largest market shares from market statistics. According to Datanyze's consolidated data, Survey Monkey and Google Surveys are the leaders in Online survey tool (Customer feedback management) market with 33.93 and 31.33 percent respectively (Datanyze, n.d.). Likewise, the leaders in quantitative analysis tool market (predictive analytics) are SPSS (56.11 percent) and SPSS statistics (11.72 percent) (Datanyze, n.d.-b).

Version	Features
Free	- Supported: Text files (pdf, docx, rtf, odt, ...), Tabular files (xlxs, xls, csv, sav, tsv, ...) - Coding, highlighting of texts - Categorizing of texts, displaying texts in groups - Simple charts, graphs of phrase frequency and coded parts - Tutorial, FAQ, basic technical support (via emails, messages)
Proprietary add-on (paid)	- Supported: Audio (mp3, flac, wav, …) - More extensive visualizations, e.g. mind mapping, word cloud, also for interrater agreement (agreement map) - Collaboration measures: • Real-time group coding of documents • Member-checking support: e.g. differentiated permission schemes/roles, sharable links, presentation for public view. - Transcribing audio files - Professional technical support (e.g. call center)
Third-party integrations	• Survey tools (Survey Monkey, Google Survey) • Quantitative analysis tools (SPSS package)

Table 6. Recommended features for two tiers of QDAcity

- **Price**

Pricing model includes two tiers in accordance with product proposition:

  - Free: as mentioned earlier, at such a level of functionalities, there are already some free options in the market – Taguette, QDAminer Lite, and other university-owned free solutions. If we are going to charge for this tier, it is will be hard for QDAcity to attract users or to differentiate itself at such a level of functionalities.
  - Premium: with a subscription for access to proprietary version and features. Since the development team has not had an estimation of total costs of operating, maintaining, further improving the product, pricing is done based on perceived value (in terms of functionalities), user's willingness to pay, and benchmarked against what competitors are charging. Certainly, the prices given by potential users are not a fixed standard for pricing. It depends a lot also on the proposition of the product as well as competition. With fewer features, it makes sense to charge a lower price for QDAcity.

    So how often should a user be charged?

    According to consolidated survey results and interpretations above, potential users tend to prefer a more predictable type of cost. Considering subscription was selected as almost frequently as one-time fee in the survey and the model itself is more suitable for cloud-based services, i.e. flexibility in demand planning, subscription of looser periods (every six months, one year) is recommended for QDAcity. This pricing allows users' clear expectation of the upfront cost in a long enough period of service. Subscription also appears more practical for users as they can choose to switch to another solution later on if the software is not up to their expectations. To keep the users' loyalty, it is also more important to focus on the provided value, i.e. enough functionalities, user experience for an affordable and flexible pricing, rather than making them stay for a sunk cost.

- **Placement**

A placement strategy that is also widely used by other competitors is to have a web-based editor where users can register for an account, log in and start using the service directly without any download or installation needed. A web-based solution allows cross-platform use which is also beneficial for the development team in terms of deployment and maintenance costs.

- **Promotion**

A potential promotion tactic is to invest in research efforts with topics revolving around qualitative data analysis methods, specifically exploring use cases with QDAcity. This will help also promote QDAcity to scholars, researchers and students who search to read about qualitative data analysis tools and approaches. If we search in Google Scholar for keyword "CAQDAS", there are currently around 16,000 results, among which more than 7,000 involving NVivo and 2,000 involving

MAXQDA. This means there has been significant effort in research of approaches in qualitative data analysis using one or more specific solutions. It is, therefore, an opportunity for the team of QDAcity to promote it. Provided there were sufficient research about QDAcity, the chance a researcher or student gets to read or learn about it and considers using it would also be higher.

Additionally, referral schemes can help attract new users through word-of-mouth. One example of such a referral scheme is when an existing user refers the solution to a friend or family member, he can receive some free months of subscription. The new referred user, if he decided to subscribe to QDAcity, can also enjoy some discount. This scheme would help bring more revenues to QDAcity which otherwise would probably not be possible. Furthermore, with the power of word-of-mouth marketing, QDAcity could reach a wider base of users faster without advertising costs.

Besides referral schemes, discounts can also help enhance revenues over time. If users have been using QDAcity for some time, e.g. at least one year, they may receive a discount for upgrading to the proprietary version. This enhances not only the chance they will start paying for the service, but also their loyalty with QDAcity, especially when they already have their data and work stored with QDAcity.

Furthermore, a highlight of unique selling points of QDAcity in website homepage, advertising and landing page would also allow QDAcity to also stand out.

Last but not least, security measures which can potentially sway potential users with security concerns to adopt a cloud-based software should also be included and emphasized in advertising, homepage and landing page. For internet users who are used to web-based solutions, they may not be so impressed with a promotion message about those measures. However, it always helps gain trust even for these users when it is clear how the service provider goes to different extents to ensure security for customers.

6.2 Market size and revenue estimation

A portion of students are also working at universities as research assistants. In calculating estimated sales and revenues, those who marked themselves as researchers but are also students will be considered only in student groups in order to avoid duplicates. In other words, researcher groups should not include anyone who is also a student, because pricing is more advantageous towards students. As per common sense, they will choose that for students rather than for researchers.

Since QDAcity upon launch would most likely be promoted first in the German scholar communities (students, partners and networks of the faculty or university), populations in Germany are used for estimations of sales in the first three years. The potential market share can still be larger considering QDAcity is a cloud-based solution and not geographically restricted.

Below are different factors that contribute to sales and revenue estimations.

- **Target population statistics**: the total number of students (Bachelor, Master), PhD candidates, researchers (industry and government) in Germany will be

considered the base population for sale and revenue estimations. The reason why no specific figures are retrieved for every scientific field, despite this being asked in the survey, is because the number of respondents in each field is too few to make meaningful extrapolation to the populations. In addition, there is skewed distribution of respondents among the fields. Thus, only estimation per user groups is done.

Figures for populations are taken from OECD statistics – graduates by year. The reason for using number of graduates instead of enrolled students is because it is not certain the enrolled ones make it through all the years up to graduation. The graduating number is much lower than the enrolled number and thus gives a more conservative estimation for overall populations.

Since the statistics for students on OECD homepage are available by the time of collection only up to 2019, an average of the 5 years from 2015 to 2019 is used to reduce impact of temporary fluctuations. The numbers are listed in Table 7.

Students	**2015**	**2016**	**2017**	**2018**	**2019**	**Average**	**Avg. cumulative growth**
Bachelor	318,662	324,315	327,687	320,295	391,905	336,572	
Growth	-	1.77%	1.04%	-2.26%	22.36%		5.31%
Master	196,668	202,984	212,850	216,373	222,689	210,312	
Growth	-	3.21%	4.86%	1.66%	2.92%		3.16%
PhD	29,218	29,303	28,404	27,838	28,690	28,690	
Growth	-	0.29%	-3.07%	-1.99%	3.06%		-0.45%
Total	544,548	556,602	568,941	564,506	643,284	**575,574**	

Table 7. Population figures and growth in each user group

Statistics for researchers on OECD homepage are available in different years for industry and government sectors. Therefore, average of each sector is calculated first and summed up together for final figure for researchers, as in Table 8.

Researchers	**2015**	**2016**	**2017**	**2018**	**2019**	**Average**	**Avg. cumulative growth**
Industry	586,030		623,125		667,394	625,516	
Growth			3.12%		3.49%		3.30%
Government	62,790	62,840	63,862	66,978	69,352	65,164	
Growth		0.08%	1.63%	4.88%	3.54%		2.52%
Total	648,820	-	686,987	-	736,746	**690,680**	

Table 8. Population figures and growth for each group of Researchers

The estimated population for researchers is therefore 690,680.

- **Penetration rate**
  Penetration rate is understood as the percentage of the relevant population that has at least used a similar product or product of similar category once within a specific period (Farris et al., 2010). To narrow down the penetrable population and give a

more conservative estimation of market size, only those who claim to be interested in purchasing a CAQDAS for personal use, not have security concerns or be deterred from inherent issues of a cloud software are considered.

Used percentages inferred from survey responses are as follows in Table 9.

	Ever used (a)	Interested and no concerns/ not deterred from adopting (b)	Penetration rate (a) x (b) (*)
Bachelor, Master	17.3%	100.0%	17.31%
PhD	32.0%	50.0%	16.00%
Researchers	50.0%	50.0%	25.00%

Table 9. Estimated penetration rates for each user group

- **Frequency of use:** Estimation for sales and revenues are limited to proportions of potential users who have from 3 projects or more per year, with each lasting at least more than three months. First, because these have more projects (equal to or more than three) per year and are more likely to justify their needs for an annual or long-term subscription. Second, a project lasting longer than three months in analysis may imply more intensive workload which more likely makes them feel worth buying a subscription.

  Table 10 lists the different combinations of answers for the two questions "In a typical year, how many projects do you have where you need to code and analyze qualitative data?" and "For each project, how long in average it takes you to complete coding and analyzing qualitative data?". Among these, as explained, we consider only the highlighted combinations in estimation of sales and revenues.

Number of projects	**Length per typical project**
One-time event	Weeks to less than 1 month
One-time event	1-3 months
One-time event	More than 3 months
1 or 2	Weeks to less than 1 month
1 or 2	1-3 months
1 or 2	More than 3 months
3 to 5	Weeks to less than 1 month
3 to 5	1-3 months
3 to 5	More than 3 months
5 or more	Weeks to less than 1 month
5 or more	1-3 months
5 or more	More than 3 months

Table 10. Survey options regarding project frequency and length

Not accounting other users who have less frequency and shorter period of project work does not mean they would not buy or generate any revenues. However, in the worst-case scenario, with limited need of use, these users may prefer a free solution.

Table 11 specifies the percentages of people with at least three projects per year with each lasting more than three months, calculated using the survey results.

Group of users	Number of projects	Length per typical project	Out of all responses for respective user group (%)	Sum per user group (%)(**)
Bachelor, Master students	3 to 5	More than 3 months	0.00%	**6.67%**
	5 or more	More than 3 months	6.67%	
PhD students	3 to 5	More than 3 months	0.00%	**0.00%**
	5 or more	More than 3 months	0.00%	
Researchers (non-students)	3 to 5	More than 3 months	0.00%	**9.09%**
	5 or more	More than 3 months	9.09%	

Table 11. Percentages of respondents with more frequent project periods

The listed percentages inferred from survey results will be used for extrapolation to user populations to arrive at the total addressable market in the later part.

- **Type of CAQDAS**

  As mentioned, QDAcity is better positioned as an "in-between" solution rather than to lean towards either a "focused, lightweight" or "feature-rich" one. Thus, if we narrow the sales of QDAcity down to only those who prefer a focused, lightweight software, that would probably underestimate the actual potential of QDAcity. The potential market of QDAcity can also be larger as users of feature-rich solutions may not need all offered functionalities and come to realize that they are overpaying for their needs. Those users may very likely seek a more affordable solution which satisfy their personal needs well enough. Considering such potential, the achieved populations will not be further narrowed down to either "feature-rich" or "focused, lightweight" preferring users, but serve as the final sales estimate.

- **Annual Recurring Revenue per user (Annual subscription fee)**

  The fee or recurring revenue is achieved through benchmarking with what competitors are offering, how many functionalities QDAcity would be providing per marketing plan and how much potential users are willing to pay.

  The reason for choosing an annual recurring revenue instead of a monthly one, as mentioned earlier, is because it simplifies the estimation of expected total revenues while fitting in with the recommended pricing model of loose-period subscription. Correspondingly, the estimations assume only payment from users who have regular and longer-spanned research projects throughout a typical year as this segment of users are more likely to purchase an annual subscription.

  Benchmarked against what other products are charging, an annual recurring revenue per user of 36 Euro (3 Euro per month – paid annually) for students and 72 Euro (6 Euro per month – paid annually) for non-students are reasonable to

assume. Considering the scope of features to be offered by QDAcity is limited compared with competitors and the premium is only for the add-on features, it is plausible to charge at around half of what others are charging.

- **Average Annual Cumulative Growth**

  One important factor in sales and revenue forecasts is average annual growth rate. Assuming no change in pricing (or average annual recurring revenue per user), an annual growth rate in revenues is also that in sales and it involves growth in two aspects:

  - Average annual growth in target populations:
    Populations of students and researchers may be subject to changes over time, due to, for example, increasing focus on economics or education. The average growth rate in target populations is estimated using historical cumulative growth of the three user groups. Derived average cumulative growth rates for every user group from "Target Populations" section are listed in Table 12 again for reference:

	Bachelor	Master	PhD	Industry Researchers	Government Researchers
Average annual growth	5.31%	3.16%	-0.45%	3.30%	2.52%
Average population	336,572	210,312	28,690	625,516	65,164

Table 12. Average annual population growth per user group

    Since there are significant deviations in populations of the three user groups, access to different segments of the target market is also not the same. Other things equal, QDAcity will be more likely to be known and/or used by a researcher than any user in the other user groups, because of its larger population. If the development team wants to target PhD students more, they will need to invest more in promoting QDAcity or reaching out to this group. Also, because of the small population of PhD students, a negative growth of this user group would have trivial impact on the future size of accessible market, compared to that of the other groups.
    Due to this reason, there is no need to normalize before averaging the growth rates. The average growth across all populations is: 2.77%

  - Average annual growth in penetration of the software over time (sales):
    Estimations are made with respect to three scenarios: "Pessimistic", "Most likely", and "Optimistic". According to benchmark information by Serena Capital, an investor and provider of operational support for ventures, European SaaS companies with annual revenues less than €1 Million grew at an annual rate from 50 percent to 260 percent, with a median of 150 percent (Serena Capital, 2020). Meanwhile, SaaS Capital, a firm that acts as a fund manager for SaaS companies, published research about growth of private SaaS businesses with data collected from over 1,400 enterprise participants. Specifically according their survey in 2020, the median and mean growth

rates of private SaaS companies are 68 percent and 144 percent respectively for those earning less than $1 Million per year (SaaS Capital, 2020). There is a considerable deviation between the two median numbers from the two surveys, probably due to the differences in scale, location, and currency in which revenues are measured. Thus, to reduce the impacts of these factors, we can use the average of 68 and 150, 109 percent, as the estimated annual growth rate for QDAcity.

This average value of the market can be a safe estimation for the "Optimistic" scenario because CAQDAS is a mostly dedicated market for researchers, academic workers and students. Even among these, some may not have the need for a CAQDAS, but rather use a more general-purpose solution for data analysis (La Pelle, 2004). Thus, CAQDAS market is expected to belong to the lower end of the growth spectrum compared to other SaaS markets. The "Pessimistic" scenario assumes 0 percent in growth (no growth at all) and "Most likely" scenario can reasonably take the mid-way value between the other two – 54.5 percent.

Considering both population growth and sales growth, the three scenarios of growth estimations are as listed in Table 13.

Pessimistic	Most likely	Optimistic
2.77%	57.27%	111.77%

Table 13. Three scenarios of growth estimation for revenues

Growth in sales and revenues is, however, offset by the following factors:

- Annual churn rate:
  This rate represents how many users (in percentage) stops using the product. Cloud-based software can be accessed from anywhere and with little effort. Therefore, there are a lot more alternative products for users and switching between different software after some time is not difficult. The churn rate for cloud-based applications, therefore, should be high, especially for start-ups and newcomers in the first several years. It requires time until they can maintain sustainable customer base with the right strategies. A total churn rate consists of two parts: voluntary churns (preferences, switching to other solutions, no need for use anymore) and involuntary churns (technical issues leading to cancellation of subscription, payment failure, etc.) (Krull, 2020).

  Due to lacking specific statistics for the market of CAQDAS, figures with respect to the industry and business size are used instead. According to Recurly Research survey running for 12 months in 2018 over a sample of 1,500 subscription site, the average total churn rate for SaaS industry B2C is 5.06 percent per month (Recurly Research, n.d.). Meanwhile, the average total churn rate for Average Monthly Revenue Per Customer of less than $10 is 7.12 percent (Recurly Research, n.d.). According to Recurly benchmarks (n.d.), businesses with higher Revenue Per Customer are also more volatile and have higher churn rates. This helps confirm again that having a monthly revenue or price of less than $10, which also aligns with interpretations from

the competitor analysis, is recommendable.

The mean of these two benchmarking values can be used as the estimated monthly churn rate for QDAcity: 6.09 percent.

Since sales are calculated per annum, an annual churn rate is needed. Churn rate is directly correlated with retention rate as: Churn rate = 1 – Retention rate. Retention rate reflects the percentage of users who still stay by the end of the month or year of use, depending on the considered period. Monthly churn rates, as derived from Recurly, are calculated after each month, which means the retained users continue to be users in the following month. Due to this reason, an annual retention or churn rate is a cumulative figure of the monthly rate over the period of twelve months.

Accordingly, the annual churn rate to be used for QDAcity can be calculated as follows:

$$1 - (1 - 6.09\%)^{12} \approx 53\%$$

Final calculations of market size and forecasts are presented in Table 14.

	Bachelor, Master	PhD	Researchers	Total
Population	546,884	28,690	690,680	1,266,254
Penetration rate (*)	17.31%	16.00%	25.00%	
Regular users (**)	6.67%	0.00%	9.09%	
Market size	6,313	0	15,696	22,009
Average ARR per user (USD)	36	36	72	
Estimated revenues (USD)	227,281	0	1,130,091	1,357,371

Table 14. Estimation of revenues for QDAcity

The forecasts are limited only to three years (with same average growth, churn rate, and prices) because further than that the business itself as well as the economic landscape may have certain changes that undermine the initial assumptions.

Based on the first year estimated revenues of 1,357,371 USD, the revenues of the next two years are as follows:

Estimations	Pessimistic	Most likely	Optimistic
Growth rate	2.77%	57.27%	111.77%
Churn rate	53.00%	53.00%	53.00%
Revenues – 2nd year (USD)	675,564	1,415,331	2,155,098
Revenues – 3rd year (USD)	336,228	1,475,765	3,421,649
Average over three years	789,721	1,416,156	2,311,373

Table 15. Revenue forecasts for QDAcity in the first three years

As calculated in Table 15, over the first three years, QDAcity is expected to achieve an average of 1,416,156 USD per annum in German market. The figures are 789,721 USD and 2,311,373 USD for the worst- and best-case scenario respectively.

As mentioned earlier, a cloud-based solution is not bounded by geographical conditions, and therefore can be used by people who pursue higher education and research activities around the globe. Assuming no difference in preferences of people in different locations, this means an opportunity for the business of QDAcity to penetrate a total population of approximately 5,300,000 for the European Union markets and 16,500,000 for up to 40 countries (38 in OCED[1]) spanning across the globe (OECD, 2021). These 40 countries include: Austria, Australia, Belgium, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States, Brazil, and Russia (OECD, 2021).

The reason for choosing OCED figures is because the statistics are publicly available for reference about specific user groups. It may not cover all countries globally but include information of those that focus on development and socio-economic initiatives. Among the countries whose population figures are available on OECD, non-OECD like Brazil and Russia are also included. These are also growing economies with increasing focus on research, education and technology (Export-Import Bank of India, 2014). For a software like QDAcity which targets specifically user groups of students, scientists, and researchers, it also makes sense to look at these 39 countries rather than a total of 195.

Specific figures by OECD in year 2019 are listed in Table 16.

	BA, MA	PhD	Researchers	Total
OCED-All	12,490,729	308,511	3,705,591	16,504,831
EU	3,444,318	93,190	1,844,999	5,382,507

Table 16. Population statistics for each user group in 2019 by OECD

(OECD, n.d.) (OECD, 2021)

6.3 Limitations

Even though the data collection, interpretation and analysis are carried out conservatively and with considerations, assumptions do put limitations to validity of the results and discussions. Such limitations can be attested with follow-up surveys and research. Specifically, they are

[1] Organisation for Economic Co-operation and Development includes countries that collaborate on developing and improving education and social policy.

- Limitations in the survey-related choices:
    - The survey does not ask about location while the populations reached include students and researchers in also some other countries in Europe and English-speaking countries, which means the results assume similarities among target users who live and work in different countries. This may not be the reality and impact the validity of sales estimations.
- Limitations inherent from data collection process:
    - Despite an effort in reaching more people in the fields with fewer responses in the first place, the survey ended up being filled by more students or researchers from Information and Communication technologies, Business administration and Law compared to others. Figure 42 shows response distribution among different scientific fields. While almost 19 percent of all respondents are working or studying in the field of Information and Communication technologies, only 8 percent of them in Education or Arts and Humanities. This unbalanced distribution may deliver a biased view about preferences and behavior of potential users in general, if there should be such differences among those of different fields.

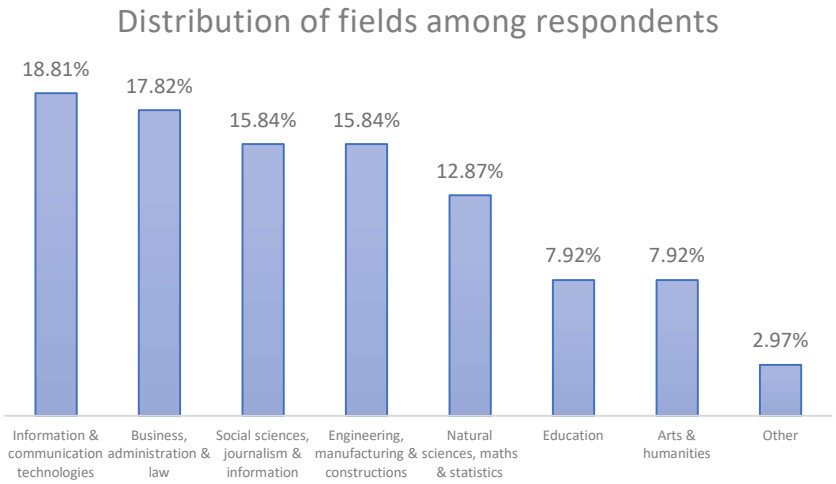

Figure 42. Distribution of scientific field among respondents

- Limitations in interpretations of survey results:

  There are certain assumptions on which the estimation of market size and revenues are based. Following questions should also be asked and further investigated to ensure no biases in interpretations and recommendations:

    - Are there potential users who actually prefer an in-between solution, neither a feature-rich nor a lightweight, focused one?
    - Is the majority of users who have more and longer projects in a year willing to pay for an annual subscription?

- Are a majority of potential users who prefer an integration between a survey tool or quantitative analysis tool with CAQDAS actually using Survey Monkey, Google Survey, and/or SPSS?
- Do a majority of potential users actually want a long-term subscription (i.e.1-year time)?

Different answers to above-listed questions would cause sales estimations and recommendations to deviate. However, the limitations should not make the interpretations or recommendations invalid, because they are based on comparison reasoning between two or more objects as well as common sense. That means it should be applicable in a majority of situations. Also, there are several criteria applied in estimation to keep the figures conservative enough, which also helps to retain the lowest possible impact of biases.

It is better nonetheless to figure out if there is any outlier to what has been expected from users' preferences and behavior, and if such an outlier is significant enough to affect or negate any of the conclusions or recommendations. The next step for this research is thus to follow up again with the same user groups to validate such points.

7 Conclusion

In conclusion, there is a substantial potential market for personal use of QDAcity in Germany. As QDAcity is a cloud-based solution and not geographically restricted, it may benefit also from a larger potential from student and researcher populations in European Union and the other 23 countries around the world. Based on specific product offers, users' behavior, and pricing, QDAcity is expected to bring about more than 1,400,000 USD every year for the first three years in Germany only. This is estimated assuming an annual recurring revenue of 36 Euro per year for students, 72 Euro for non-students, a two-tier product offer: a free one and a proprietary add-on with additional features.

Using empirical data collection, this thesis delivers insights about CAQDAS potential market, users' preferences and behavior, which almost none of the existing literature has addressed before. Specifically, besides basic features which QDAcity already has, there is also an observed preference amongst potential users of the software for processing of tabular files, audio files; collaboration features: real-time group coding, interrater agreement, member checking; visualization; transcribing and professional support. Backup, autosave measures, basic support and tutorials are without question important for almost any cloud-based software. They also proved to be important among survey respondents or potential users of CAQDAS. There is also additional demand noticed for third-party integrations with survey and quantitative data analysis tools. As for pricing, users tend to be indifferent between subscription and one-time pricing models despite their tendency to quantify costs in a lump sum.

From the survey results, there are also hints into some other demanded features and expectations of users for a CAQDAS, for example "automated coding", memo-ing, multilingual support, multiple user pricing, processing of JSON and image formats. These can also be the topic for further investigation regarding specific implementation and actual demand among CAQDAS users, if development team has an intention to include them in QDAcity in the future.

From all these findings, the thesis presents business recommendations that facilitate determination of product proposition, pricing, placement and promotion for QDAcity.

It is important to note that all recommendations are vulnerable to deviation in effectiveness. Since the data collection, analyses and estimations are based on assumptions, these can fall prey to biases and limitations as mentioned earlier, and thus may not be that effectively applicable if any of the assumptions no longer holds. Thus, before the full-scaled launch of QDAcity, there suggests follow-up studies or a Beta launch to validate interpretations against assumptions included in this thesis.

Appendix A

The datasets extracted from Lime survey tool and all used aggregation results are included in a workbook that comes in a digital version together with this thesis paper.

As mentioned in chapter “Research Method” (chapter 4), below is the list of Facebook groups of students and researchers where the survey was shared.

Vietnamesische Studenten in Deutschland

Vietnamese PhD mums

Internationals of Netherlands

International Students in Denmark

International students in Norway

International Students UK

International Students in France

International Students in Australia (ISA)

International Students in Canada

International students in Finland

The Research Survey Exchange Group

PhD Postdoc in US - UK - AU and Canada

Doctoral Research, PhD students, Bursary, Scholarship & Jobs, UK

UX Designers & UX Design

UX Researchers Association

Research Participation - Dissertation, Thesis, PhD, Survey Sharing

FAU International Master Computational & Medical Engineering

FAU Erlangen: Suche/Biete

List of universities contacted for fill-in of surveys:

University of Trier, Germany

University of Freiburg, Germany

University of Bonn, Germany

Technical University Dortmund, Germany

University of Konstanz, Germany

University of Augsburg, Germany

University of Stuttgart, Germany

University of Kassel, Germany

University of Amsterdam, the Netherlands

University of Cambridge, the United Kingdom

References

ATLAS.ti. (n.d.). ATLAS.Ti Homepage. Retrieved November 5, 2021, from https://atlasti.com/

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member Checking. *Qualitative Health Research*, *26*(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Costa, A. P., Reis, L. P., Souza, F. N. D., & Moreira, A. (2017). *Computer Supported Qualitative Research: Second International Symposium on Qualitative Research (ISQR 2017) (Advances in Intelligent Systems and Computing)* (1st ed. 2018 ed.). Springer.

Dalkin, S., Forster, N., Hodgson, P., Lhussier, M., & Carr, S. M. (2020). Using computer assisted qualitative data analysis software (CAQDAS; NVivo) to assist in the complex process of realist theory generation, refinement and testing. *International Journal of Social Research Methodology*, *24*(1), 123–134. https://doi.org/10.1080/13645579.2020.1803528

Datanyze. (n.d.-a). *Customer Feedback Management Software market share report*. Retrieved November 2, 2021, from https://www.datanyze.com/market-share/customer-feedback-management--27

Datanyze. (n.d.-b). *Predictive Analytics Software Market Share report*. Retrieved November 2, 2021, from https://www.datanyze.com/market-share/predictive-analytics--366

Dedoose Features. (n.d.). Dedoose. Retrieved November 5, 2021, from https://www.dedoose.com/home/features

Evers, J. (2018). Current Issues in Qualitative Data Analysis Software (QDAS): A User and Developer Perspective. *The Qualitative Report*. Published. https://doi.org/10.46743/2160-3715/2018.3205

Export-Import Bank of India. (2014, November). *Research & Development In Brics: An Insight* (No. 168). http://www.brics-ibcm.org/images/publications/Research_and_Development_in_BRICS-An_Insight.pdf

Farris, P., Bendle, N., Pfeifer, P., & Reibstein, D. (2010). *[Marketing Metrics: The Definitive Guide to Measuring Marketing Performance By Farris, Paul (Author) Hardcover 2010]* (2nd ed.). Pearson Education, Inc.

Freitas, F., Ribeiro, J., Brandão, C., Reis, L. P., Souza, F. N., & Costa, A. P. (2017). *Learn by yourself: The Self-Learning Tools for Qualitative Analysis Software Packages*. Digital Education Review.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, *9*(3), 330–338. https://doi.org/10.1016/j.sapharm.2012.04.004

Glaser, B. G., & Strauss, A. (1967). *The Discovery of grounded theory*. Van Haren Publishing.

Global Qualitative Data Analysis Software Market Size By Product Type, By Material, By Geographic Scope And Forecast. (2021, April). Verified Market Research. https://www.verifiedmarketresearch.com/product/qualitative-data-analysis-software-market/

Hackett, P., & Hayre, C. M. (2020). *Handbook of Ethnography in Healthcare Research*. Routledge.

Jew, S. H. (2014). *Some Types of Data Visualizations in NVivo*. Scalar. https://scalar.usc.edu/works/using-nvivo-an-unofficial-and-unauthorized-primer/types-of-data-visualizations-in-nvivo

Jones, M. L. (2007). Using Software to Analyse Qualitative Data. *Malaysian Journal of Qualitative Research*, *1*(1), 64–76.

Kalpokaite, N., & Radivojevic, I. (2019). Teaching qualitative data analysis software online: a comparison of face-to-face and e-learning ATLAS.ti courses. *International Journal of Research & Method in Education*, *43*(3), 296–310. https://doi.org/10.1080/1743727x.2019.1687666

Kawulich, B. (2004, January). *Qualitative Data Analysis Techniques*. RC33 (ISA), Amsterdam, the Netherlands. https://www.researchgate.net/publication/258110388_Qualitative_Data_Analysis_Techniques

Krull, A. (2020, April). *Subscriber Retention and Understanding Involuntary vs. Voluntary Churn*. Recurly Research. Retrieved November 2, 2021, from https://recurly.com/blog/subscriber-retention-and-understanding-involuntary-vs.-voluntary-churn/

Kuckartz, U. (2019). Qualitative Text Analysis: A Systematic Approach. *ICME-13 Monographs*, 181–197. https://doi.org/10.1007/978-3-030-15636-7_8

La Pelle, N. (2004). Simplifying Qualitative Data Analysis Using General Purpose Software Tools. *Field Methods*, *16*(1), 85–108. https://doi.org/10.1177/1525822x03259227

MaxQDA. (n.d.). MaxQDA Homepage. Retrieved November 5, 2021, from https://www.maxqda.com/

NVivo. (n.d.). QSR International. Retrieved November 5, 2021, from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

O'Kane, P., Smith, A., & Lerman, M. P. (2019). Building Transparency and Trustworthiness in Inductive Research Through Computer-Aided Qualitative Data Analysis Software. *Organizational Research Methods*, *24*(1), 104–139. https://doi.org/10.1177/1094428119865016

OECD. (2021). *Research and development (R&D) - Researchers - OECD Data*. The OECD. Retrieved November 2, 2021, from https://data.oecd.org/rd/researchers.htm

OECD. (n.d.). *OECD Graduate by Field*. The OECD. Retrieved November 2, 2021, from https://stats.oecd.org/Index.aspx?DataSetCode=EDU_GRAD_FIELD

QDAMiner Features. (n.d.). Provalis Research. Retrieved November 5, 2021, from https://www.google.com/search?q=qdaminer+6&oq=qdaminer+6&aqs=chrome.0.69i59j0i13j0i22i30l2.1788j0j4&sourceid=chrome&ie=UTF-8

QDAMiner Lite Version. (n.d.). Provalis Research. Retrieved November 5, 2021, from https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/

Recurly Research. (n.d.). *Churn rate benchmarks*. https://recurly.com/research/churn-rate-benchmarks/

Robert M. Schindler (1991) ,"Symbolic Meanings of a Price Ending", in NA - Advances in Consumer Research Volume 18, eds. Rebecca H. Holman and Michael R. Solomon, Provo, UT : Association for Consumer Research, Pages: 794-801.

Serena Capital. (2020, February). *2020 European SaaS Benchmark*. https://european-champions.org/wp-content/uploads/2021/02/2020-european-saas-benchmark-by-serena-1.pdf

Silver, C., & Lewins, A. (2014). *Using Software in Qualitative Research: A Step-by-Step Guide* (Second ed.). SAGE Publications Ltd.

Taguette Home. (n.d.). Taguette. Retrieved November 5, 2021, from https://www.taguette.org/

Verdugo-Castro, S., Sánchez-Gómez, M. C., García-Holgado, A., & Costa, A. P. (2019). Mixed methods and visual representation of data with CAQDAS. *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality*. Published. https://doi.org/10.1145/3362789.3362847

Woods, M., Paulus, T., Atkins, D. P., & Macklin, R. (2016). Advancing Qualitative Research Using Qualitative Data Analysis Software (QDAS)? Reviewing Potential Versus Practice in Published Studies using ATLAS.ti and NVivo, 1994–2013. *Social Science Computer Review*, *34*(5), 597–617. https://doi.org/10.1177/0894439315596311