MSc Information Systems Development

Masters Dissertation

Title

The design and evaluation of a system for the analysis of wellbeing data

Submitted in partial fulfilment of
the requirements of Edinburgh Napier University
for the Degree of Master of Science in
Information Systems Development

Author

Christian Heinz Schwab
40217437

School of Computing
August 2016

Authorship Declaration

I, Christian Heinz Schwab, confirm that this dissertation and the work presented in it are my own achievement. Where I have consulted the published work of others this is always clearly attributed. Where I have quoted from the work of others the source is always given. With the exception of such quotations this dissertation is entirely my own work. I have acknowledged all main sources of help. I have read and understand the penalties associated with Academic Misconduct.

I also confirm that I have obtained **informed consent** from all people I have involved in the work in this dissertation following the School's ethical guidelines

Signed:

Date: 15th of August 2016

Matriculation no: 40217437

Data Protection Declaration

Under the 1998 Data Protection Act, The University cannot disclose your grade to an unauthorised person. However, other students benefit from studying dissertations that have their grades attached.

Please sign your name below *one* of the options below to state your preference.

The University may make this dissertation, with indicative grade, available to others.

The University may make this dissertation available to others, but the grade may not be disclosed.

The University may not make this dissertation available to others.

Abstract

The aim of this work was to design a system for the analysis of wellbeing data. It follows a previous study conducted with six nurses in which their heart rate variability was measured and which serves as an indicator for the stress level and the body resilience against these stressors. The recorded data of the study was evaluated to assess the recording reliability of the used Polar H7 instrument, and to raise questions about the data for further research. Based on that, the requirements on which the system was developed were defined. Furthermore, the designed system was evaluated by three experts in the field of healthcare through semi-structured interviews and questionnaires in order to ensure the system included the key functionalities needed to benefit further research. For the implementation of the system, Python – a multipurpose programming language with a broad offer of libraries in data science, was investigated. The operational capability of the key parts of the system could be proved, showing Python's suitability for the further development of it. Results showed that the strengths of the system are to combine multiple recorded data elements within a screen and enable its users to detect reasons for outliers and inconsistencies in the recorded data by using video material. Furthermore, the data over multiple subjects can be compared and evaluated. As discussed during the evaluation process of the data, differences in the heart rate variability of the participants could be detected. Nevertheless, research is still ongoing, and there are not conclusive results yet. However, a foundation was built with the system in this work, to further analyse the data, find patterns and answer the arisen questions.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to sincerely thank my supervisor Alistair Lawson, for his patient guidance and helpful feedback during the entire dissertation project. Furthermore, I want to thank Dr. Lu Fan for his useful advice at the beginning of this project.

Especially, I would like to thank my parents and my brother who were always there for me, not only during this work, but throughout my entire life. I also want to thank my friends and Nuria, who provided me with useful critiques and advice during the whole year as a student.

Vielen Dank

1 Introduction

The health care sector is known for offering physical demanding occupations which are often correlated to stress (Kakunje, 2011). A nurse has to offer a high level of care, which requires a decent fitness level (Amtmann & Amtmann, 2010). A good fitness level helps therefore to effectively perform the required activities for the nursing endeavours such as patient transfer, transport of equipment or other activities that require physical performance (Amtmann & Amtmann, 2010). Nevertheless, there exist few studies about the fitness state of nurses (Albert, Butler & Sorell, 2014).

Nursing is a challenging profession which includes not only high professional performance prospects, but also a high extent of public performance expectation. Furthermore, long varying working hours and high level of stress are also associated with this profession (Schluter, Turner, Huntington, Bain & McClure, 2011; Videman, Ojajarvi, Riihimaki & Troup, 2005). The resulting mental stress and physical exposure in this job are therefore triggers for health issues and physical wear over time (McNeely, 2005).

A survey was completed by 1452 hospitals in the United Kingdom to evaluate the wellbeing and health of their staff, whereby a big part were nurses. The study showed that 43% of the participants were overweight, and had a low understanding of the importance of physical activity or even acknowledged healthy diet existed (Blake, Mo, Lee, & Batt, 2012). This correlates to the findings with a study carried out by Zapka et. al, (2009) about 194 nurses of six different hospitals, who were mostly overweight and less committed to physical activities or a healthy diet.

Several factors such as long shifts or challenging working environments with different patients, may therefore influence the adoption of an unhealthy lifestyle and cause different health issues like obesity (Blake et. al, 2012). In regard to this, often stated reasons for low physical engagement or an unhealthy diet are less free time or tiredness after work (Blake et. al, 2012). Nevertheless, it is important to keep in mind that the human resources are crucial for the healthcare sector, and the wellbeing of the nurses is essential for a high level of care (Bryant, 2011). Publications of the World Bank have mentioned midwives and nurses as resources which are very cost efficient for providing a high standard of quality care in the healthcare sector (Hongoro & McPake, 2004). Following this, human resources are a key part of the health system but have been a less attended factor in the development of the health-system over time (Hongoro & McPake, 2004).

Information technology has the potential to help improve nursing practices and efficient care (Huntington et. al, 2009; Gelsema et. al, 2006). This includes the use of technology to analyse data to detect or minimise stress levels and improve the physical state of nurses (Yuan et. al, 2009; Edwards & Burnard, 2003). This Masters dissertation project therefore aims to design a system for integrating, visualizing, aligning, cleaning and analysing already collected wellbeing data from a research study which investigated the performance of nurses during several activities completed in a period of one hour.

The study was conducted with six nurses separated in two cohorts, each of three individuals, and carried out on two days (each day one cohort). The study setup included nine different tasks of 2 to 4 minutes each. Over one hour, these tasks were continuously repeated by each participant from two to four times.

The study used the Polar H7 heart rate sensor to collect the heart beat to beat interval, the beats per minute, speed, distance, acceleration and cadence of each participant. To be better able to monitor the participants, videos were recorded for three tasks, showing the participants' movements during the tasks. Apart from that, Tanita scales were used to measure each participant's health factors (Weight, BMI, Body fat etc.), and questionnaires were conducted to collect the socio-demographics of the participants.

Therefore, this work will concentrate on designing a system to deal with these different types of data and enable users to efficiently analyse the heart rate variability of the participants; which can be seen as an indicator for the stress and exhaustion level of an individual (Sweetwater, 2011; Tsuji et. al, 1996).
As this work aims to conduct research on how a beneficial software for the analysis of wellbeing data looks like; the research takes an information systems development stance rather completely adopting a software engineering approach.

For the implementation of the system, the focus is set on Python, a multi-purpose programming language which has several strengths in regard to data analysis and visualisation (McKinney, 2013; Mehlhase, 2014). Furthermore, this work deals with three research questions which are going to be answered throughout the project.

- RQ1: What questions might be asked of the data?

- RQ2: Does the equipment (Polar H7) ensure a reliable data recording?
- RQ3: To what extent does the software tool enable the research to draw conclusions about the data and provide a foundation for further investigations in this area?

To meet these aims, the work will start with a literature review in Chapter 2; which provides the theoretical background about the heart rate variability data recorded in this project as well as the equipment used in the study. Furthermore, the state-of-the art in Data research and analysis as well as the user acceptance in regard to systems and technology are discussed. Last but not least, a theoretical foundation about Python is provided for the later implementation parts of the system. Chapter 3 presents a review of the methodologies used in this work; followed by a discussion of the requirements of the software tool to be developed and an evaluation of the recorded data of the study in Chapter 4. On this basis, the design of the prototype system will be presented in Chapter 5. It will also include an overview of its users, the evaluation of the prototype through interviews with experts in the field of wellbeing and healthcare, and a discussion of the software architecture. Chapter 6 comprises the implementation and testing of the requirements of the system in Python, presenting the proof of concept of this work, followed by the evaluation of the entire work in Chapter 7 and final discussion in Chapter 8.

2 Literature review

The literature in this work was collected through tertiary methods which included search engines and indexes which conducted the author to secondary literature sources as journals articles, conference proceedings or handbooks (Saunders & Lewis, 2009). The rationale of the collected literature is to frame a theoretical foundation for the objectives and research questions of this thesis.

2.1 Analysis of wellbeing data

The following two chapters discuss the heart rate variability investigated in this work, as well as the equipment used to record the different data.

2.1.1 Heart rate variability and cardiovascular stress

Heart rate variability has been increasingly investigated in the areas of cardiovascular disease, stress and work out in the last decades (Boullosa et. al, 2012; Buchheit et. al, 2010; Uusitalo et. al, 2011; Yarnell, 2008). Stress in general can be described as a reaction to a specific activity or situation where certain mental and (or) physical requirements are imposed on a person (Sharma, 2005). In general, small quantities of stress can have empowering effects on people whereas enormous amounts of stress can have a negative influence on the health and performance of human beings (Sharma, 2005). Stress is a relevant topic especially in healthcare due to the sometimes fatal and irreversible consequences of errors (Familoni, 2008).
The administration of stress is done by the autonomic nervous system which can be divided into sympathetic nervous system – which is the driver for thrilling states in the human body; and the parasympathetic nervous system, which is the indicator for balance and emotional self-regulation (Park et. al, 2009; Berntson & Cacioppo, 2004). In the case when an individual experiences stressors (physical or mental), the autonomic nervous system gets activated, the sympathetic nervous system is triggered, and the parasympathetic system is held back (Akselrod et al., 1981).
This process can be related to the inner body reactions in a stress scenario: norepinephrine and epinephrine hormones are shot into the blood; which results in adequate reactions of the body as for instance muscle tensions, blood vessels or increased blood pressure. This further leads to a different heart rate (beats per minute)

and to a decrease in the heart rate variability (Taelman, Vandeput, Spaepen, & Huffel, 2009).

Therefore, the interactions of the sympathetic and parasympathetic nervous system have a high effect on the R-R Interval, which can be described as the time-related fluctuation between two heart beats in the electrocardiogram (ECG). The time-related fluctuations between the different R-R intervals is defined as heart rate variability (HRV) (Billman, 2011; Berntson et. al, 1997; Sweetwater, 2011).

Figure 1: R-R interval (Aubert et. al, 2003)

Clinical research about the HRV indicates that high heart rate variability levels are related to a stronger physical resilience of the body and lower experienced levels of stress. Therefore, low heart rate variability levels are an indicator of higher experienced body stress levels and a lower body resilience (Sweetwater, 2011; Tsuji et. al, 1996). Thus there exist various analysis methods for HRV as the time-domain, frequency-domain and other nonlinear methods (Aubert et. al, 2003; Tsuji et. al, 1996; Uusitalo et. al, 2011; Chandola et. al, 2008).

Nevertheless, to accurately measure stress levels with the R-R interval, the interval has to be converted from the time-domain into the frequency-domain; where a low and a high frequency can be determined, and a ratio between low frequency over high-frequency is an indicator for the experienced stress level of an individual (Konold et. al, 2011; Sweetwater, 2011).

However a myriad of factors have an influence over the heart rate variability; for instance emotional circumstances, physical fitness levels, respiration and inspiration, age or cardiovascular deseases (Boullosa et. al, 2012; Taelman et. al, 2009; Tonello et. al, 2014; Sweetwater, 2011; Williamon et. al, 2013). Because of this, a range for a normal heart rate variability is regarded as difficult to asess because of its great dependency on various factors and its resulting dynamism, so every individual is seen to have his own range (Sweetwater, 2013).

2.1.2 Equipment used in the study

In the completed study, Tanita scales were used to analyse body measures as visceral fat, bone mass, metabolic age, muscle mass, body water, body fat or weight of the participants (Tanita, 2016). Furthermore, the Polar H7 heart rate sensor strap was used. This tool records heart rate data for real-time heart monitoring (bpm and R-R interval), speed, acceleration and running cadence. Within the sensor strap, a bluetooth functionality is implemented which allows to connect to a fitness app or another training instrument (Polar, 2016).

Nevertheless, data "measurement errors" (Lopes & White, 2006, p.54) could appear with the sensor strap due to various reasons (Gamelin et. al, 2008; Giles et. al, 2016; Haynes & Lisicki, 1996; Lopes & White, 2006; Polar, 2016):

- Movement artefacts which are caused through sudden impulsive actions during activities (including coughing, sneezing, abrupt movements etc.)
- Drop outs, where no event is detected within more than 2 seconds (R-R interval)
- Poor skin contact
- Electrical interference due to the so called "sixty-cycle noise" (Haynes & Lisicki, 1996, p. 1215) from a close electrical source, which could be light, motors, cables, monitors amongst others.

But how can users deal with such errors during the analysis of the data?
Literature indicates that error values have to be detected and corrected to ensure a high data quality for further processing of the data through time-or frequency analysis methods, which would lead to interferences in the results without a manipulation of the errors (Berntson & Stowell, 1998; Mateo & Laguna, 2003; Solem et. al, 2006).

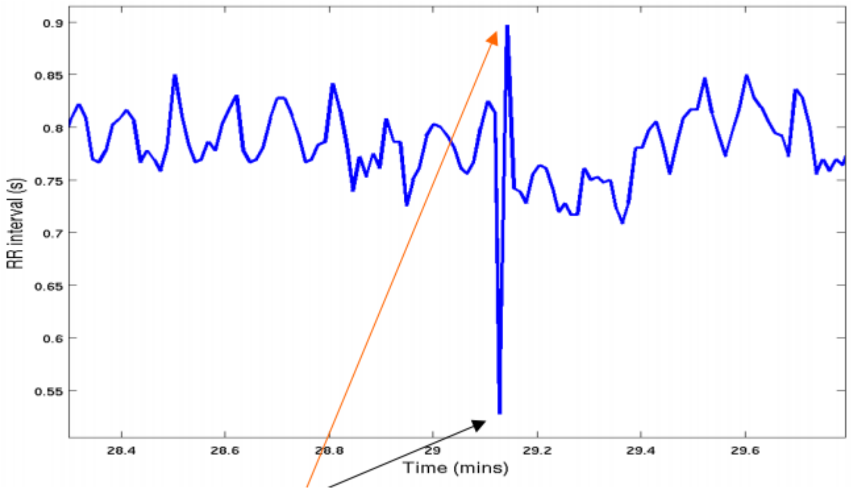


Figure 2: Example for an outlier in the interval (Clifford, 2006)

Literature indicates there exist different ways to deal with these outliers. Different authors mention a modification of these values in regard to the other previously accepted R-R intervals (Kemper, Hamilton, & Atkinson, 2007; Berntson & Stowell, 1998; Mateo & Laguna, 2003). However, the removal of outlier intervals “performs as well as or better than more complicated methods for these relatively short data samples” (Lippman, Stein, & Lerman, 1994, p. 411), and therefore represents an appropriate option to clean the recorded R-R interval from outliers (Altini, 2016; Kamath, Watanabe, & Upton, 2013; Buchheit, 2014; Karlsson et. al, 2012). Therefore, Kamath et. al, (2013) suggests to remove all R-R intervals which differ more than 200 milliseconds (ms) from the preceding one. Different authors suggest to exclude R-R intervals which differ by more than 20% (Altini, 2016; Pipilis, Flather, Ormerod, & Sleight, 1991; Kleiger, Miller, Krone, & Bigger, 1990). Minor aggressive cleaning methods of the R-R interval are also mentioned, for instance, the cleaning of intervals where the difference to the preceding one is above 50% (Karlsson et. al, 2012; Altini, 2016).

2.2 Data Science

The term Data Science represents actions that aim to detect underlying correlations between the data and further build data models within applications around the data to be investigated (McClure, 2015; Dhar, 2012; Provost & Fawcett, 2013). Data Science is the extraction of knowledge from data through intensive processing with the aid of tools in information technology and theories from mathematics and statistics (Dhar, 2012). Nevertheless, in comparison to big data it doesn’t focus on the quantity of data, but rather on the approach; using different data mining techniques, data-analytic thinking, and visualisation techniques for a target-oriented knowledge retrieval (Provost & Fawcett, 2013; McClure, 2015). However, gathering, using and processing information is often a challenge due to sensitive information and privacy regulations of the data that have to be taken into account (Kambatla et. al, 2014; Intel IT Center & Intel, 2012).

2.2.1 Big Data

Each day, 2.5 exabytes of data are created from various applications and tools. This reveals the world's proficiency for information creation which has never been that powerful and excessive since information technology arose with the start of the 19th century (Wu et. al, 2014; McAfee & Brynjolfsson, 2012). The rapid increase in the importance of information is affecting every industrial sector from small to large sized organisations. All of them deal with enormous amounts of data resulting by continuous technical advancements in device features which enable the digitalisation and recording of various contents (Hey et. al, 2009; Villars & Olofson, 2011). Followed by that, the tag "BIG DATA" evolved with time, representing high volumes of data, containing a tremendous amount of information that is created by people about people or processes, presenting opportunities for advancing data science in different sectors (Boyd & Crawford, 2012; Hampton et. al, 2013).

Accordingly, when talking about Big Data there exist in literature various ways on how to specify this term (Assunção et. al, 2015). The majority of the authors (Lalani & Meniya, 2015; Zikopoulos & Eaton, 2011; Trnka, 2014; Sagiroglu & Sinanc, 2013) use the terms *volume*, v*ariety* and *velocity* as the three key elements which determine Big Data. In addition to that, *value* as well as *veracity* as other factors of Big Data are widely common in literature (Assunção et. al, 2015; Fan & Bifet, 2013; Kaisler et. al, 2013; Chen et. al, 2014; Swan, 2015). Therefore the 5V that describe Big Data can be described as (Chen et. al, 2014; Trnka, 2014; Fan & Bifet, 2013; Kaisler et. al, 2013):

Table 1: 5V of Big Data

5 V of Big Data	Characteristics
Volume:	Represents the amount of data that is steadily increasing
Variety:	Describes the different data types existing, which are *unstructured*, *structured* and *semi-structured* data
Velocity:	Refers to the continuous streams of content and the speed of its generation and frequency with which it is delivered
Value:	Describes the benefit which is generated through the data
Veracity:	Refers to the data reliability and trustworthiness

A lot of this data which fits the characteristics of the five V can be found outside of data pools due to the high prevalence of the World Wide Web. Furthermore, technical instruments as for example sensors or mobile devices open opportunities for the analysis of the data using appropriate toolsets to identify patterns and draw conclusions from it (Villars & Olofson, 2011; Fan & Bifet, 2013; Hampton et. al, 2013). This suggests that data is more than promising not only for business applications but also in multiple areas like the financial industry (transactions), the social media domain or also the healthcare and human welfare sector which represent industries with a largely growing data size (Chen et. al, 2014; Kambatla et. al, 2014). Institutions therefore face challenges in generating and managing big datasets (Khan et. al, 2014). The difficulty of this is not only to gather and access data of a high volume, velocity and variety; but also taking into account the variability of the data and retrieve valuable insights from it (Bakshi, 2012).

In this work, the recorded data of the six participants are no Big Data yet, but the focus is to develop a system which is scalable and has the potential to deal with more participants and thus larger amounts of data.

2.2.2 Data types

Applications don't just generate large amounts of data, data also include different types, which are commonly divided in literature as structured data, unstructured data and semi-structured data (Assunção et. al, 2015; Wu et. al, 2014; Fan & Bifet, 2013;

Baars & Kemper, 2008; Khan et. al, 2014; Goli-Malekabadi et. al, 2016; Sagiroglu & Sinanc, 2013).

Structured data can be described as data allocated to fields and are often integrated in tables and relational databases (Baars & Kemper, 2008; Russom, 2011). *Unstructured data* represent content that does not include a defined data model and cannot easily be stored in relational tables as textual content, videos, images or audio files (Zikopoulos & Eaton, 2011; Bakshi, 2012). *Semi-structured data* include all data that are intertwined with the Extensive Mark-up Language (XML) or have a similar structure as web pages or social media posts, are not predefined to fields but include tags so the different data entities can be differentiated (Baars & Kemper, 2008; Russom, 2011).

Nevertheless, most of the generated data are unstructured or semi-structured, so the combined use of structured and unstructured data becomes an important factor to make the right decisions and operate efficiently (Vedder et. al, 1999; Baars & Kemper, 2008). This work deals with structured data (spreadsheets) and unstructured data (video data) and aims to integrate and combine both data types to be better able to analyse the recorded data and draw conclusions from them.

2.2.3 Time-series data

In a lot of scientific projects, measurements and data acquisitions are carried out over a defined timeframe and generate data which are time-based, thus called time series data (Esling & Agon, 2012, Wu et. al, 2005; Liao, 2005). Time series data are often generated in industrial activities, patient treatments or through measurement of different metrics (Wu et. al, 2005). In time-series data, periodic fluctuations are frequent especially in medical and economic time series, whereas also non-periodic structures as trends or correlations exist. Nevertheless, the internal structure of time series data helps to process the data though querying and visualising to receive valuable insights from the them (Wu et. al, 2005; Esling & Agon, 2012; Fu, 2011).

The modelling of time series data with the help of applications and toolsets is not only helpful especially when forecasting or monitoring data, but also generates a better understanding of the data content and is therefore common in data science and research (Wu et. al, 2005). This work will therefore deal with time-series data recorded by the Polar H7 strap sensor which contains the R-R interval of the participants and other HR, Speed/Distance measurements (see appendix 1.3).

2.3 Data analysis strategy

As discussed in section 2.2, data created by scientific tools as sensors, mobile and web applications are increasing and become an important factor in many areas (Hey et. al, 2009). Therefore, the capability to analyse data and retrieve value from the analysis is crucial in today's society and depends on the integration of distributed content and computational practices (Luckow et. al, 2015). This indicates strong requirements for new tools and methods with the capability to help users in monitoring different kinds of data (Fayyad et. al, 1996).

The term Knowledge discovery in databases (KDD) describes the detection of useful and unrecognised insights from data through a variety of practices such as natural language processing, distributed programming, statistical and visual analysis, pattern recognition, data mining, sentiment analysis and human computer interaction (Sagiroglu & Sinanc, 2013; Jagtap & Kodge, 2013). However, an efficient data analysis process normally requires different steps to be completed as the selection of data, sanitisation and anonymization of data, data transformation and integration, cleaning and pre-processing, data mining and data visualisation. (Cios et. al, 2010; Fayyad et. al, 1996; Soibelman & Kim, 2002; Jagtap & Kodge, 2013; Liu, 1996).

2.3.1 Selection of data

The first step of a successful data analysis process is questioning what the aim of the analysis process is and what issues the analysis process takes on (Trnka, 2014). Based upon this, the relevant data have to be retrieved from the database and loaded into a table, spreadsheet or into a tool to further process and analyse them (Jagtap & Kodge, 2013; Varian, 2014).

2.3.2 Data sanitisation and anonymization

Privacy is unneglectable for modern life (Bishop et. al, 2010; Fung et. al, 2010). As already mentioned in the previous chapter, data science has to consider privacy regulations and sensitive information when analysing data. Oliveira and Zaane (2003) stated that Atallah et. al, (1999) were one of the first authors who discussed sanitisation of data as a process of altering sensitive data in a way that they have a less identifiable value for the user.

Researchers use different terms to describe the disguising of sensitive data and also use different methods and theories on how to protect sensitive data especially in databases (Bishop et. al, 2010; Dasseni, Verykios, Elmagarmid, & Bertino, 2001; Koukis, Antonatos, Antoniades, Markatos, & Trimintzios, 2006; Nelson, 2015; Oliveira & Zaane, 2003). Bishop et. al, (2010) for example state, that sanitisation as well as anonymisation are both stated in the literature; and describes anonymisation as a way to protect the privacy of a specific entity, whereas sanitisation can be seen as a hypernym for data in general. Nelson (2015) for instance divides the disguising process in de-identification, which represents the method of removing any content that is personally identifiable and aims on lowering the risk of unwanted disclosure of information about the identity, whereas anonymisation refers specifically to individual entities as well. Oliveira and Zaane (2003) describe the act of disguising information as sanitisation process to clear “restrictive association rules that contain sensitive knowledge” (Oliveira and Zaane, 2003, p. 613).

But what is actually determined as being sensitive data? Individual privacy regulations determine sensitive data as personably identifiable information (PII) (Bishop et. al, 2010; Shmatikov & Narayanan, 2010). Personally identifiable information is regarded as content in different data types which allows the observer to relate it to an individual (Shmatikov & Narayanan, 2010) and includes various factors as (McCallister, Grance, & Kent, 2010) :

- Personal characteristics (photographic images, especially faces)
- Address data (including email address)
- Personal identification details (Name or also credit card number, passport number, security number etc.)
- Other information about an individual which can be linked to the individual (for instance race, religion, employment information, medical information, education information, financial information)

Therefore, especially raw data in natural shape typically consists of sensitive content which is revealing information about individuals. This data has to be carefully dealt with to respect the individual privacy especially in regard to analysing the data (Atallah et. al, 1999; Bishop et. al, 2010; Fung et. al, 2010; Nelson, 2015). With a look at the data existing in this work, only the video deals with personably identifiable information in regard to the faces of the participants, which can be recognised in the recording.

2.3.3 Data transformation and integration

Data transformation is sometimes needed to integrate the content in a tool. It can therefore be seen as the altering of the data structure or representation of the data before they are integreated into a tool (Rahm & Do, 2000). Following this process, the term data integration describes the combination of several sources of data (Jagtap & Kodge, 2013). Especially when dealing with different types of data as for instance structured and unstructured information, integration and aligning of the data is an efficient and straight forward way to give the user simultaneous access to the data sources for further analysis, which is mostly done by a graphical user interface in a software tool (Baars & Kemper, 2008; Becker et. al, 2002; Priebe et. al, 2003).

2.3.4 Cleaning and pre-processing of data

Data cleaning and pre-processing is an action which aims at improving the quality of the content by removing inconsistencies or larger errors in the data sources (Rahm & Do, 2000), correct or eradicate duplicated records or missing information (Agusthiyar & Narashiman, 2014). Normally different tools as for instance OpenRefine, Weka or Data Wrangler, but also MS Excel can be used for cleaning and pre-processing the data especially in regard to large quantities, to minimize a manual review of the data and unnecessary programming effort (Varian, 2014; Jagtap & Kodge, 2013).

2.3.5 Data mining

The term data mining represents “the analysis of observational datasets to find unsuspected relationships and summarise the data in novel ways that are both understandable and useful to the data owners” (Hand et. al, 2001, p. 1). Data mining aims therefore at two points (Kohavi, 2000):

Insight: Detects patterns and tendencies in the data which are understandable, can be individualised and enable users to make decisions based on the findings.

Prediction: Is used to model the data and predict events based on the raw data, for instance the affinity of customers to buy a certain product based on their buying habits in the internet or their demographic data.

The data mining approach takes into account that the detected value of the data is unrecognised, not superficial and useful for the spectator (Fayyad et. al, 1996). The analysis of the data is often seen as a process where either supervised or unsupervised learning techniques are applied. Supervised learning can be divided in decision trees, rule algorithms, general statistical methods and neuronal networks; whereas unsupervised learning can be parted in association rule algorithms and clustering methods (Cios et al., 2010). Within these methods the two major rules when looking at data mining algorithms are efficiency and scalability. They have to be scalable and efficient in order to deal with large amounts of data and aim to optimize the output result (Jiawei & Pei, 2011).

2.3.6 Visualisation and interpretation of data

Visualisation aims to use the visual capabilities of human beings to provide cognition with the help of computer-created representations that are scalable and implementable on a large set of data. Especially in science and data analysis, visualisation is a key step to get a better understanding of the data sets and their contained value (Mann et. al, 2002). The visualisation of data aims to summarise the data and make its content more understandable to the user with the help of charts, tables and graphical acquisitions. Data is visualised in various cases to enable the discovery of patterns or associations in the data, which also leads to further questioning of the data and especially facilitates decision making processes whereas a non-visualisation approach may complicate it to detect such findings and may also decelerate decisons (Jiawei & Pei, 2011; Berry & Linoff, 2004).

Especially within this work visual analytics is emphasized. Visual analytics aims to lower complex cognitive tasks allowing individuals to directly interact with the data, providing a better understanding about them and further enabling to draw conclusions and make decisions (Keim, Mansmann, Schneidewind, Thomas, & Ziegler, 2008). Overall, visual analytics can be described as "Information visualization [which] is about developing insights from collected data" (Fekete, Wijk, Stasko, & North, 2008, p.3). In regard to this work, visualisation of the data is helpful with respect to two points (Fakete et. al, 2008):

- Provides a framework for the interpretation and analysis of time series data.
- Establishes opportunities to integrate and synthesize different types of information from different sources into a unified data representation so users get deeper insights about the data and are enabled to draw conclusions from them.

2.4 Technology Acceptance Model

This chapter will outline theories which have to be taken into account when designing the software tool in a later step in this work. Within the development of information technology, technical and quality aspects of a tool are crucial, but also several other factors regarding its future users have to be taken into consideration (Davis, 1989; Hsu et. al, 2009; Padilla-Meléndez et. al, 2013; Umarji & Seaman, 2005; Wallace & Sheetz, 2014).

Social science mentions different theories in literature whereby the Technology Acceptance Model is the most frequently taken approach to assess users' affinity to a technology (Kim & Chang, 2007; Yarbrough & Smith, 2007).

The Technology Acceptance Model (TAM) based on Davis' work (1989), is an information system theory which highlights the *perceived ease of use* and the *perceived usefulness* as the two major factors in regard to the user's acceptance of information technology; and has had fundamental empirical and theoretical support in literature over years (Padilla-Meléndez et. al, 2013). The perceived ease of use expresses the complexity and the level of physical and mental user-involvement that an information system acquires (Davis, 1989). The perceived usefulness can be defined as the "degree to which an individual believes that using a particular system would enhance his or her job performance" (Davis,1993, p. 477).

The perceived usefulness as well as the perceived usability of a system are influenced by several external factors as education level, gender or the person's affinity to information technology (Burton-Jones & Hubona, 2005; Venkatesh et. al, 2012). Nevertheless, also different other factors appear in literature as the perceived *attractiveness* of the system or the *individual attitude* of the user which influence the use of information systems (Umarji & Seaman, 2005). Apart from that, also Rogers (2010) with his "uncertainty reduction process" (p.216) states different factors (trial ability, relative advantage, complexity, observability or compatibility) during an innovation process which help to reduce uncertainty about a technical innovation and predict the adoption of its future users.

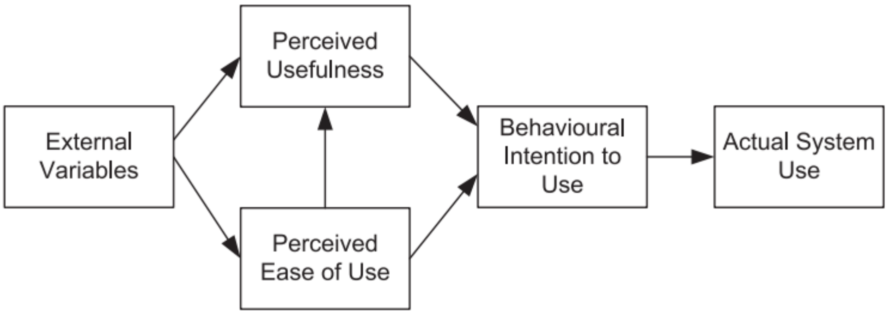


Figure 3: Technology Acceptance Model (Brezavšček, Šparl, & Žnidaršič, 2014)

These theories have to be taken into account within the development process of the software in this project to increase the user acceptance and intention of use of the system to be developed. Especially through interviews in a later step of this work, the innovation tool is evaluated to adjust and improve it.

2.5 Python for Data Analysis

Python, an open source, multi-purpose scripting language stands for fast code development and modification, with no need for specific development tools as for example demanded in Java (McKinney, 2013; Mehlhase, 2014; Laura et. al, 2013; Java, 2016). For developers, Python provides a fast development environment which facilitates an easy workflow and efficient prototyping also due to the fact that Python scripts are uncompiled and additionally allows for the integration of other programming languages such as C or C++ for instance (Laura et. al, 2013; McKinney, 2013).

Another benefit is the big active science community of Python, offering several free available libraries especially in regard to the generation of visualisations, data analysis and exploratory computing; which makes of Python a rival in regard to other programming languages and statistic tools in data science as for example Matlab, R or Stata (McKinney, 2013; Mehlhase, 2014; Perez & Granger, 2007; Laura et. al, 2013).

As described by McKinney (2013), Pandas, NumPy, SciPy and Matplotlib can be seen as essential libraries for data analysis in Python and many other existing libraries are built on: The python library Pandas provides the framework for a fast and easy processing of structured data through a combination of array-computing features enabling fast data manipulation in relational databases or spreadsheets. As McKinney

(2013, p.4) explains, it is “one of the critical ingredients enabling Python to be a powerful and productive data analysis environment.”

As discussed in the sections above, data analysis often requires a data visualisation to summarise the data and make it more accessible to the user. Therefore, Python offers various possibilities to the user for a visualisation of the data. One of the most popular libraries which enables for generating plots and various other 2D results is matplotlib. Especially in regard to its various modules it is seen as a beneficial environment for visualising and exploring data. Furthermore, it provides an interactive environment which allows the user to zoom in the graph and also save the plot (Hunter, 2007). The Numerical Python (NumPy) library includes different tools accompanying mathematical algebra functions to multidimensional arrays. Especially numpy arrays provide for numerical data effective options for manipulating and storing data in comparison to other Python data structures (McKinney, 2013; Van Der Walt, Colbert, & Varoquaux, 2011). The Scipy library builds on the calculation and computational ability existing within Numpy, offering different advanced tools for optimizing, special functions or media processing (f.e. images). One of the tools existing in Scipy is for instance a Fast-Fourier-Transform package; which allows to calculate the frequency domains of the R-R intervals in order to receive numerical stress indices for an individual besides the time-domain measures (see 2.1.1).

In addition to these libraries, other libraries and hundreds of third-party modules exist, providing an environment for data analysis with different aids and possibilities to python developers to create software (Laura et. al, 2013; Perez & Granger, 2007; Laura et. al, 2013; Mehlhase, 2014; Perez & Granger, 2007).
Therefore, in regard to the existing video data in this project; besides spreadsheets, various options for media processing exist among the many free available libraries in Python as openCV or the moviepy package which are common for editing video material (Bradski, 2000).

Furthermore, Python also provides support for graphical user interface toolkits which allow to build data-centric applications (McKinney, 2013; Perez & Granger, 2007). Different toolkits for writing graphical applications exist from the standard TkInter GUI package of Python to various third party cross-platform and platform-specific python based solutions; which include among others, frameworks as Qt, GTK, WxWidgets and Cocoa (Python, 2016; Perez & Granger, 2007). The developer can use Python’s flexibility and therefore choose the toolkit which is more suitable for his project (Perez

& Granger, 2007). Concerning the tool to be developed in this project, the possibility to build a web application should be also briefly mentioned in regard to a possible future scale up of the system to be developed. Here Python offers different possibilities as for instance Django, which is a python high level framework that can be used as a backend tool to set up web based applications for the visualisation and analysis of different data (McKinney, 2013; Rodriguez-Martinez, Seguel, & Greer, 2010).

2.6 Conclusion

Different literature sources have been reviewed to provide a theoretical foundation for the following chapters. The literature shows that the heart rate variability can be seen as an indicator for the experienced stress level of an individual. High heart rate variability can be seen in general as an indicator for a greater body resilience against stressors and accordingly lower experienced levels of stress than vice versa (Sweetwater, 2011; Tsuji et. al, 1996). In the study, Tanita Scales were used to record different health factors of the individuals. The Polar H7 sensor strap was used to record the investigated heart rate variability of the participants in this work and other heart rate, speed and distance measurements during the recording period. In literature there are various reasons mentioned for appearing errors (outliers) in the recording results of the sensor strap which have to be detected and corrected for the further analysis of the data to receive valid results (Berntson & Stowell, 1998; Mateo & Laguna, 2003; Solem et. al, 2006).

The recorded data in this work can be divided in structured and unstructured data, which are the spreadsheets and the video. As discussed in literature, the combined use of these different data types is seen as beneficial for data analysis in general and also benefits this work; as will be shown in the following chapters. Furthermore, the literature indicates different strategies to follow when analysing data. For this, Python shows different strengths in literature. One of its core strengths is that its scripts are uncompiled and can therefore be easily changed or extended with a simple text editor. Furthermore, the language provides possibilities to develop desktop as well as web applications and provides various free available libraries to the user to process various types of data. In regard to the design and development of the software system in this work, theories as the Technology Acceptance Model should be taken into account to increase the user acceptance and intention of use.

3 Methodology

This chapter gives an overview of the research methodology used for this work. The research philosophy, framework and strategy of this work are described, followed by an overview of the research tools used for data collection. Furthermore, the analysis of the obtained data as well as testing methods for the implementation part of this work are discussed.

3.1 Research Philosophy

From an ontological perspective, the work followed two approaches. An objective approach was adopted in regard to the evaluation of the raw data and the reliability of the polar instrument used, as those factors representing noticeable facts for attaining objective knowledge (Saunders, Lewis & Thornhill, 2009). However, in contrast to this, a subjective approach is adopted for the design of the software tool which will include social actors and their subjective opinions and perceptions to be integrated in the development process of the software (Saunders et. al, 2009).

Furthermore, two different epistemological approaches are adopted as well for this research. For the evaluation of the data and the reliability of the instrument, a positivist approach will be adopted. Saunders and Lewis (2012) state that the main goal of researchers following a positivism procedure is to examine observable aspects and their cause and effect. Despite a positivist approach often goes together with predictions of results and the testing of hypotheses, this work won't focus on creating and testing of hypotheses due to the small population in this study (Saunders & Lewis, 2012). It will rather generate questions about the data and design a system that allows for a more flexible analysis of the data and enables users to answer arising questions about the data in further research.

On the other side, when it comes to the design of the system, an interpretivist position is adapted, which also fits to the subjective research approach for this part of the work (Saunders et. al, 2009). Interpretivist philosophy immerses on the social context where behaviours are generated without imposing any external causes on it (Saunders et. al, 2009). The social context of this position is existing within the chosen tools for data collection in this work. These are Interviews and questionnaires to ensure a useful and efficient development of the system through user evaluations (Saunders et. al, 2009). In addition to that, the general approach of induction is chosen for this work, which focuses on exploring the collected data to evaluate the themes or issues existing and use them to design a robust system that is capable to efficiently deal with the data.

3.2 Research Framework

Since the focus of this work is on the analysis of the reliability of the Polar instrument used in the study and to develop a system which is able to answer the arising questions about the existing data; the overall research style is exploratory which "involves the investigation of a particular contemporary topic within its real-life context, using multiple sources of evidence" (Saunders & Lewis, 2012, p. 116).

In addition to that, for the development of a software tool in Python a "build" research method is followed, which includes a requirement, a design, an implementation and a testing stage (Amaral et. al, 2011). After the requirements are formulated, the system is designed in Axure, an application for fast prototyping and user interfaces (Daliot, 2013). An evaluation stage is included after the design stage with the intended user group[1] of the system. This is important to adapt required functional changes to the system and ensure a beneficial development of it (Martin, 2003).

3.3 Research Strategy

The work integrates combined methods (Creswell, 2013), using "quantitative and qualitative data collection techniques" (Saunders et. al, 2009). A qualitative research approach was conducted in form of semi-structured interviews with its user group to evaluate the design of the software tool. The questions in the interview are based on the literature of the Technology Acceptance Model (see chapter 2.4) and the research questions of this work. After the interviews, questionnaires with similar questions were given to the interviewed participants to prove the repeatability of their answers given in the interviews (Saunders et. al, 2009; Creswell, 2013).

Quantitative research methods in form of descriptive statistics were conducted as well to primarily analyse the reliability of the data recording instruments (Polar H7) and apply the selected time-domain analysis method (see section 4.2.4) for analysing the HRV to show how the data can be analysed more efficiently later with the help of the system (Saunders et. al, 2009; Creswell, 2013). In this work the focus was set on the time-domain method due to the different types of data existent in this project, which can be better aligned through a time-domain based display and processing of the data

[1] The people who are interviewed represent the user group of the later system (researchers)

(especially R-R interval and video, see section 4.2.4). Based on a comparison of different studies in the literature dealing with heart rate variability, several descriptive time-domain measures were selected for this work (see section 4.1).

3.4 Interviews and Questionnaires

This work includes interviews and questionnaires for the evaluation of the prototype design of the system. Therefore, the sampling strategy as well as the data collection and analysis methods will be discussed in the following sections.

3.4.1 Sample Selection

This work used a *purposive sampling* strategy for the evaluation of the software design in the project (Saunders et. al, 2009). This means that in regard to the complexity of the research context, a small sample of people was chosen. These are dealing with wellbeing data and therefore adjust better to the insights the research intends to gain in regard to the development of a robust tool to analyse heart rate variability. The sampling strategy within the purposive approach can be seen as heterogeneous due to the aim of collecting, describing and explaining the key themes of the outcome of various interviews. According to Patton (2002) the patterns that emerged during the interviews are the key themes which are of particular interest for the research.

3.4.2 Data collection and Analysis

Before the interviews and questionnaires were conducted, a brief explanation had been given to the interviewee about the research of this work, which provides a better understanding and confidence to the participant when answering the questions (Saunders et. al, 2009).

Interviews are discussions between two individuals or more (Kahn & Cannell, 1957) and can help the researcher to get reliable and valid information that benefits his research and project objectives (Saunders et. al, 2009). Interviews in this work were conducted with 3 experts in the field of nursing, wellbeing and healthcare; in order to evaluate the software design and ensure the development of a helpful and efficient analysis tool. Therefore, semi-structured interview questions were chosen where the researcher has a list of themes to be discussed in a given context (Saunders et. al, 2009). The questions used in the interviews are formulated openly and in a neutral tone to help to avoid bias of the Interviewee (Easterby-Smith et. al, 2008). Grummitt

(1980) states that semi-structured open questions are especially efficient in encouraging interviewees to provide detailed answer so the researcher can obtain valuable facts. Saunders et. al, (2009) explains that open questions often start with the words how, why or what and fit especially the exploratory research approach of this work. Therefore 10 questions were prepared in regards to the literature of the TAM and the research questions of this work, which should reveal beneficial information for the development of the tool. The interviewees where asked each of the questions without having a predefined order. The answers of the interview questions were audio-recorded and transcribed to control bias and receive accurate data for a further analysis (Easterby-Smith et. al, 2008). After the transcription of the interviews a coding approach was conducted to identify patterns and main themes between the participants; which can be summarised to categories and further presented in the main body of the work in a narrative analysis approach (Saunders et. al, 2009).

Apart from that, questionnaires were used as a back-up tool after each interview to test the repeatability of the participant's given answers in the interview (Saunders et. al, 2009; Creswell, 2013). For the questionnaires, closed questions are chosen underpinned by a rating scale to determine underlying attitudes of the participant (Saunders et. al, 2009). According to Navarro and Schrauf (2005), who state that the adoption of an existing rating scale is a valid method, the rating scale of Sauro and Lewis (2012) for the evaluation of a system was adopted for the questionnaires in this work. It contains a denominated "Likert scale" to measure attitudes on a 7-point ordinal scale to evaluate the degree to which the interviewee disagrees or agrees with a certain statement (Sullivan & Artino, 2013). The bipolar scales in the questionnaire are described with "strongly agree" and "strongly disagree "and represent the respondent's attitude towards the statement. The analysis of the retrieved answers from the questionnaires is in regard to the small population size conducted in a table and is analysed through a qualitative narrative method rather than a quantitative analysis approach (Saunders et. al, 2009).

Furthermore, the literature mentions that a piloting of the collection tools is important as well to ensure the asked questions generate useful responses (Saunders et. al, 2009). Therefore, the interview questions and questionnaires were pre-tested with two people. The result was that the wording of three questions had to be changed due to hesitation in answering and slight misunderstanding.

3.5 Acceptance Testing of Implementations

Acceptance testing was chosen as the applied testing method for the implementation part of this work (see chapter 6). This approach is seen as appropriate in regard to the various requirements of the system to be developed (see 4.3). According to Hoffer et. al, (2005), acceptance testing can be seen as the testing of a tool in its operating environment, to check its requirements are satisfied. The implementation part implements parts of the requirements and serves as the proof of concept of this work whereas the testing ensures that errors in the code are detected and solved to fulfil the demanded requirements (Graham et. al, 2008).

3.6 Conclusion

The following work will contain a combined method approach including qualitative and quantitative methods to answer the research questions in this work and further benefit to the design of a system which integrates all different types of data recorded; making them accessible to the user for further analysis purposes. The quantitative part of the research in this work focuses on a positivist epistemological approach for the evaluation of the raw data and includes the time-domain method for the analysis of the R-R interval. Furthermore, questionnaires are being conducted for the evaluation of the software tool design to check the reliability of the statements given by the participants of the interviews. From a qualitative standpoint, interviews were chosen for the discussion and evaluation of the designed analysis system as they are seen as highly advantageous to obtain data where questions are open-ended or complex, as it is the case in this work (Easterby-Smith et. al, 2008; Saunders et. al, 2009).

The testing of the implementation is done through Acceptance Testing to ensure that errors in the code are detected and the functionality is as defined by the requirements.

4 Requirements Analysis

This chapter summarises other studies conducted in literature and their different approaches taken to analyse heart rate variability. Based on the recorded data, the time-domain method was chosen for analysing the heart rate variability in this work. In addition to this, the raw data including the R-R interval as well as the other measures recorded by the Polar H7 were evaluated to determine the reliability of the instrument. Furthermore, the chosen time-domain analysis method was applied to three participants in this study as a proof of concept, which primarily raised questions about the data and further confirmed the need for a future system to help analysing them. Therefore, the last section lists all requirements for the system that were identified based on the evaluation as well as on previous project meetings.

4.1 Analysis of HRV in Literature

There exist different studies in literature which deal with the analysis of heart rate variability under various conditions:

Authors	Sample	Aim	Activity (A) Rest (R) Both (B)	Analysis method	Measures	Study result
Järvelin-Pasanen et al. (2012)	51 adults	Investigation in HRV of nurses with normal and extended work shifts (36h) (Different psycho-physiological strains)	B	Time-domain	Mean(RR), SD(RR), RMSSD(RR)	- Older subjects had a lower HRV indicatig an increased sympathetic activation - HRV parameters differed between work and leisure time
				Frequency-domain	LF, HF	
McNarry & Lewis (2012)	80 adults	HRV was recorded during three 6 min bouts of exercise, under moderate and havy exercise and with a 6 min break of unloaded cycling	A	Time-domain	SD(RR), RMSSD(RR)	- HRV parameters can be reliably determined during exercise - Importance of standardizing exercise intensity with regard to fitness levels if HRV is to be reliably determined
				Frequency-domain	LF, HF	
Dekker et al. (2000)	14672 adults	Investigation in the predictive value of HRV for Chronical Heart Deseases (CHD)and death from 2-minute rhythm strips	R	Time-domain	SD(RR), SDSD(RR), RMSSD(RR), pNN50(RR)	- Low HRV was associated with increased risk of chronical heart deseases and death from several causes - Low HRV is a marker of less favorable health
Taelman et al. (2009)	28 adults	The aim of this study was to examine the influence of mental stress on HR and HRV and compare between subjects	B	Time-domain	Mean(RR), SD(RR), pNN50(RR)	- Short term HRV was reduced with a mental task - Increased sympathovagal (LF/HF) balance may have been imposed.
				Frequency-domain	LF, HF	

Maciel et al. (1986)	23 adults	A Bycycle ergometer test on untrained individuals was conducted to investigate in the reactions of the sympathetic and parasympathetic nervous systems to different exercise conditions.	A	Time-domain	HR	- Increase in sympathetic activity and HR especially for higher levels of exercising and concluded lower HRV.
Ito et al. (2001)	10 adults	Investigation in HRV of nurses in the days of day-shift and night-shift under different activities (24h ECG recording).	B	Time-domain	Mean(RR)	- The cardiac autonomic activity is largely modified by the level of physical activity regardless of the clock hour, which may have clinical implications when studying the circadian fluctuations of the onset of cardiovascular disease in shift workers.
				Frequency-domain	LF, HF	
Von Amselvoort et al. (2000)	135 adults	Investigation on the effect of work-related stressors on cardiovascular autonomic regulation during 24h ECG recording.	B	Time-domain	SD (RR)	- Shift workers had lower R-R interval standard deviations than day workers during rest (sleep) time explaining higer cardiovascular desease risk - Also smokers had lower R-R standard deviations than non smoker which might be a sign for a less favorable health.

Figure 4: Different studies that are investigating in the HRV of adults – own display

Figure 2 shows that studies with different sample sizes, aims, conditions and analysis methods have been conducted to investigate the heart rate variability of individuals. Many of the authors associated lower levels of HRV to various factors as a less favourable health level, diseases, bad living habits, more intense exercise levels or shift work. The result was a generalised increased sympathetic activity in the human body of these participants during the studies and a resulting lower HRV (see also section 2.1.1). With a look at the analysis methods used in the studies, two primary methods for the analysis of HRV can be detected: the time domain and the frequency domain method. In this work, the focus is set on the time-domain method due to the aim of aligning the different types of data for the analysis process. This can be done most effectively by using the time-domain especially when plotting the R-R intervals and analysing them in comparison with the video material (see section 4.2.4).

With regard to the literature and in comparison with the time-domain measures used by other studies (see Figure 4; Aubert et. al, 2003), the following time-domain measures of the R-R interval have been selected for this work (Altini, 2016; Aubert et. al, 2003):

<u>Mean R-R interval length over the whole recording period:</u>

$$\text{Mean (RR)} = \frac{1}{n}\sum_{k=1}^{n} \text{RR}_k$$

<u>Standard deviation of the R-R intervals over the whole recording period:</u>

$$\text{SD (RR)} = \frac{1}{n}\sum_{k=1}^{n} (\text{RR}_k - Average(R - R))^2$$

<u>The square root of the mean squared differences between consecutive R-R intervals over the whole recording period:</u>

$$\text{RMSSD (RR)} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n-1} (\text{RR}_{k+1} - \text{RR}_k)^2}$$

<u>The parameter pNN50 shows the percentage of the amount of heart beats which are separated by an interval greater than 50ms over all R-R intervals:</u>

$$pNN50 = \frac{\sum_{k=1}^{n-1}((\text{RR}_{k+1} - \text{RR}_k) > 50ms)}{n}$$

Higher parameters of SD, RMSSD and pNN50 are an indicator of a greater heart rate variability and stronger parasympathetic regulation activities in the body. They can be seen as indicators for balance and self-regulation (von Borell et. al, 2007; Aubert et. al, 2003; Dekker, 2000; Taelman et. al, 2009). In addition to this, the length of the R-R interval is influenced by inspiration (short R-R intervals) and expiration (long R-R intervals), which give insights about the breathing activity of an individual (EKG Interpretations, 2016).

The goal of this work, as already mentioned in the introduction, is the design of a prototype system for users to access and analyse the recorded data. The focus is therefore not set on the specific analysis of the data to set up hypothesis and draw conclusions. However, this work applied the selected time-domain analysis method outlined in 4.2.4 to indicate how the heart rate variability of an individual can be analysed and which questions arise when monitoring the data.

4.2 Raw Data Analysis

In this chapter, an overview of the raw data records is given, followed by an examination of the recorded data by the sensor strap for errors and missing values. Furthermore, the raw R-R interval is evaluated for all the individuals, showing outliers in the data which have to be cleaned before the time-domain measures can be applied (see section 2.1.1). The last part will outline the resulting requirements whose definition is based on previous project meetings as well as on the findings discovered through the monitoring the raw data.

4.2.1 Existing data

The data of this project was recorded from six nurses. The Polar sensor recorded for each participant the R-R interval (see appendix 1.2). In addition to that, the recording time, the daytime, the heart rate (beats per minute), the speed (km/h), the distance, the acceleration (m/s^2) as well as the running cadence were also recorded. The data were stored in a text file from a virtual cloud (R-R interval) as well as in an excel file (HR data and speed, distance measurements) (see appendix 1.3). Furthermore, videos for the tasks 3 (Simulated chest compression) 4 (assist in transferring a person from a bed to a chair) and 5 (Tying someone's shoelace) exist for each participant, showing their movements during a specific task. In addition to that, there is a file showing the different health factors (weight, body fat, BMI etc.) and socio-demographic characteristics (age, living habits, etc.) of each participant; as well as their answers about the usability of the equipment used (see appendix 1.1). In addition, another excel file exists which contains the timetable of the activities, including the time of the day and the task completed by each participant (see appendix 1.4).

4.2.2 Errors and Missing Data

With a first look at the data alongside the R-R interval recorded by the Polar instrument, an occurrence of missing values and errors can be quickly detected, which is found in the data file of each participant:

	A	B	C	D	E	F	G
1	Time	Time of day	HR [bpm]	Speed [km/h]	Distance [m]	Acceleration [m/s²]	Running cadence
2	0:00:00	10:40:39					
3	0:00:01	10:40:40	131	0.0		0.0	0
4	0:00:02	10:40:41	135	1.0	0.0	0.0	0
5	0:00:03	10:40:42	135	4.6	0.0	0.0	30
6	0:00:04	10:40:43	136	4.4	1.3	0.0	59
7	0:00:05	10:40:44	136	3.9	2.4	0.0	57
8	0:00:06	10:40:45	137	3.7	3.5	0.0	54
9	0:00:07	10:40:46	137	4.1	4.5	0.0	58
10	0:00:08	10:40:47	141	4.1	5.7	0.0	57
11	0:00:09	10:40:48	141	3.9	6.8	0.0	57
12	0:00:10	10:40:49	141	0.8	7.8	0.0	57
13	0:00:11	10:40:50	142	0.6	7.8	0.0	0
14	0:00:12	10:40:51	143	0.0	8.2	0.0	0
15	0:00:13	10:40:52	143	0.0	8.2	0.0	0
16	0:00:14	10:40:53	138	0.0	8.2	0.0	0
17	0:00:15	10:40:54	138	0.0	8.2	0.0	0
18	0:00:16	10:40:55	138	0.0	8.2	0.0	0
19	0:00:17	10:40:56	138	0.0	8.2	0.0	0
20	0:00:18	10:40:57	138	0.0	8.2	0.0	0
21	0:00:19	10:40:58		0.0	8.2	0.0	0
22	0:00:20	10:40:59		0.0	8.2	0.0	0
23	0:00:21	10:41:00		0.0	8.2	0.0	0
24	0:00:22	10:41:01		0.0	8.2	0.0	0
25	0:00:23	10:41:02		0.0	8.2	0.0	0
26	0:00:24	10:41:03		0.0	8.2	0.0	0

Figure 5: Sample data of Nurse 1

In all the files of the six participants "errors" like those from Figure 5 can be detected. Therefore, the table below shows in red a summary of the number of seconds where no parameter value was recorded for each participant (also as a percentage from the overall recording time).

Table 2: Evaluation of the recording reliability of the polar instrument

Participant	Overall time of recording	HR	Speed	Distance	Acceleration	Running cadence	Task - time recording
Participant 1	17220s	78s (0.45%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	0s (100%)
Participant 2	17293s	32s (0.19%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	0s (100%)
Participant 3	17224s	12s (0.07%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	0s (100%)
Participant 4	6344s	46s (0.73%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	12s (0.07%)
Participant 5	6062s	20s (0.33%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	0s (100%)
Participant 6	6361s	9s (0.14%)	1s (0.01%)	2s (0.01%)	Error (0 value also with speed change)	1s (0.01%)	0s (100%)

Normally each second of the recording should display a recorded parameter value (see Figure 5). What can be said through looking at the six data files is, that mostly in the beginning and in the end of the recording, the instrument didn't record every parameter value. Literature displays different reasons for this phenomenon, as for instance, a low contact between skin and the electrode, as well as possible dry skin or movement artefacts (Gamelin et al., 2008; Giles et al., 2016). However, table 2 shows that besides the continuous 0 values for acceleration which are present in every recording of each participant, the missing parameter values are minor. However, in literature there are no guidelines which state how many missing values are "normal"; due to the various possible reasons for appearing inconsistencies in the data recordings using a sensor strap (see section 2.1.2).

Especially during the recording time of the tasks which is most important, there were nearly no parameter values missing; apart from a 5 second and a 7 second interval without a heart rate value from a participant. For the time of the recording where the missing values are present, there exists no video material in this project showing the participants actions which might have provided insights about the reasons for the poor recording in that time period. Therefore, it cannot be precisely determined what the reasons for the missing values are. However, a poor skin contact, especially in the beginning (when attaching it to the chest) and in the end of the recording time (when taking down the sensor); seem to be one logical reason for this kind of data errors in the population of this study (Gamelin et. al, 2008; Giles et. al, 2016). Nevertheless, the continuous zero value for the acceleration parameter cannot be explained in this work. However, it seems to be a software error due to its recurrent appearance in every recorded file.

4.2.3 R-R Interval Evaluation

Apart from the above mentioned parameter values, the R-R interval from each participant was recorded with the sensor strap. Figure 6 shows the recorded R-R interval for Nurse 1. The intervals for Nurse 2 to 6 can be found in appendix 2.

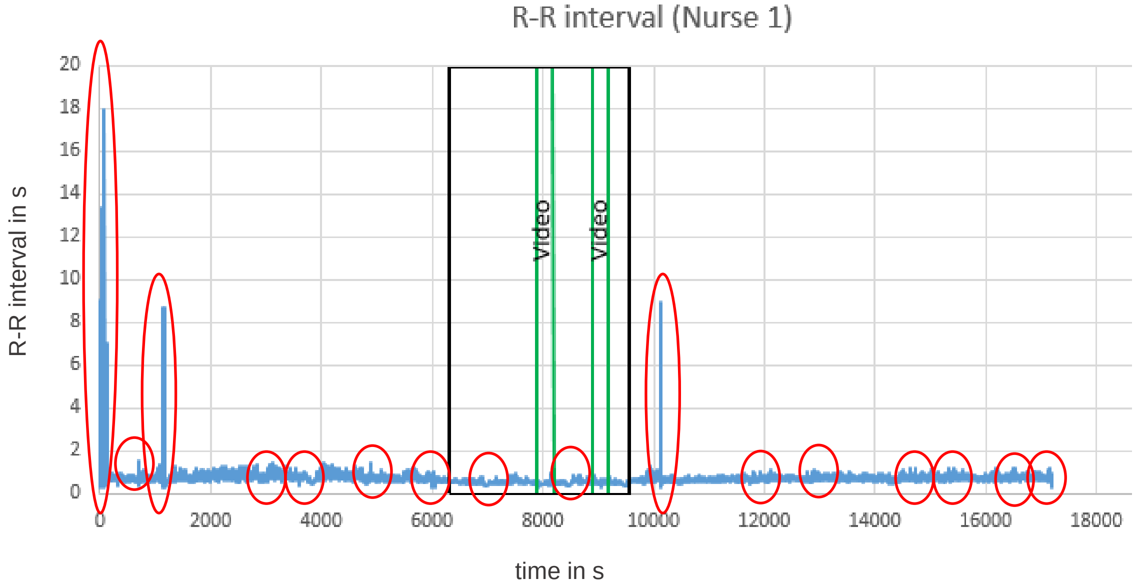


Figure 6: Nurse 1 – R-R interval

The black box in the graph represents the period of time the task recording took place, while the green lines display the period of time the video material of the participant is available (Participants wear the chest strap before and after the task recording period). With a look at the diagram, strong outliers can be detected (marked with a red circle) where no heart beat was recorded by the instrument for several seconds. As discussed in literature (see section 2.1.2), these arrhythmic intervals have to be removed in order to ensure a high quality of the data for further data processing methods (see appendix 3.1 for an example).

However, through looking at the whole R-R interval as displayed in Figure 6, not all outliers can be detected all of a sudden. With zooming at shorter time periods of the R-R interval, further outliers can be found, as shown in the examples below:

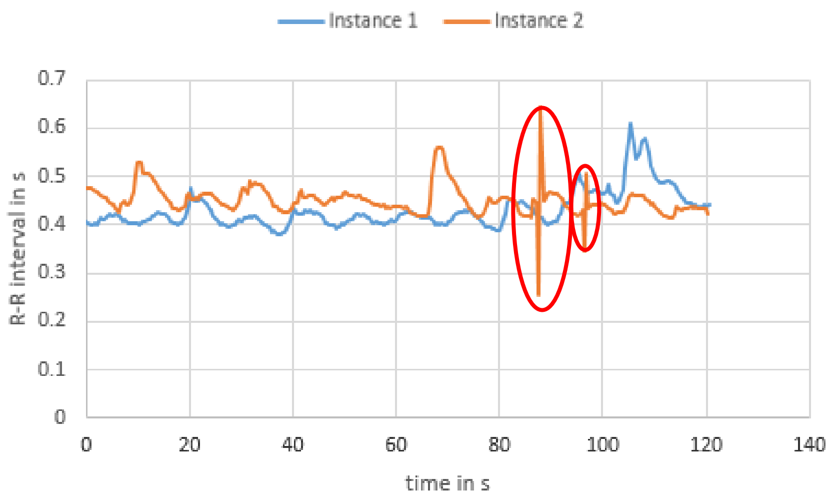


Figure 7: Nurse 1 - Task 3

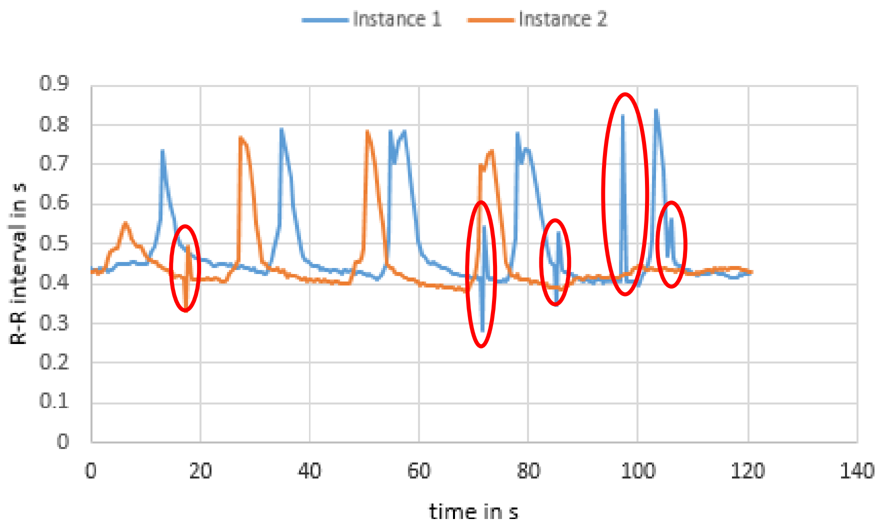


Figure 8: Nurse 2 - Task 3

Figures 8 and 9 both show the recorded R-R interval of Nurse 1 and Nurse 2 for task 3 with its two instances. In both task instances, outliers can be detected which have to be removed.

As proposed in section 2.1.2, authors suggest the cleaning of intervals which differ more than 200 milliseconds from the preceding one, whereas other methods exist where intervals differentiating more than 20% or more than 50% should be removed (Altini, 2016; Pipilis, Flather, Ormerod, & Sleight, 1991; Kleiger, Miller, Krone, & Bigger, 1990; Karlsson et. al, 2012).

Therefore, with a look over all the recorded R-R interval data, different amounts of outliers can be detected among the participants, dependent on the chosen cleaning parameter (see appendix 3.2). The results show that the number of outliers recorded by the instrument is overall low with 3.33% in comparison with the whole recorded values. However, as discussed in section 4.2.2, in literature there are no guidelines which state how many outliers are “normal” due to the various possible reasons for appearing inconsistencies in the data.

R-R intervals which differentiate more than 20% from their preceding value showed the highest numbers among the cleaning parameters, which implies a more detailed detection of outliers. With a look at the six participants, participant 1 and participant 4 differentiate from the other participants in that their data contain a higher number of outliers for each cleaning parameter. This difference raises questions about the reasons of this phenomenon. One explanation could be the stated reasons for outliers in section 2.1.2. However, further research has to be conducted with more participants to be better able to detect patterns and draw conclusions from the results.

With a look back to the outlier examples of task 3 shown in Figure 7 and 8 for Nurse 1 and 2, the findings in appendix 3.3 strengthen the aforementioned assumption that a cleaning approach for intervals which differentiate more than 20% shows the best results (Altini, 2016; Pipilis et. al, 1991; Kleiger et. al, 1990).

4.2.4 Time-domain method

With regard to the applied time-domain method in this work, the heart rate variability of the participants can be visualised and statistically analysed. Especially in combination with the video data – which allows for the visual analysis of movement artefacts and stress and exhaustion levels; the time-domain method was investigated further in this work to combine the analysis of the video and the plotted heart rate variability graph of a participant during a certain task instance.

With regard to appendix 1.4, each participant performs 9 tasks and repeats each task at least one time during the recording period. But in order to get a participant’s R-R interval of a task instance to determine the heart rate variability during that time period, an algorithm needs to be used. Figure 9 shows the algorithm which has to be developed in order to receive the R-R interval for a task instance:

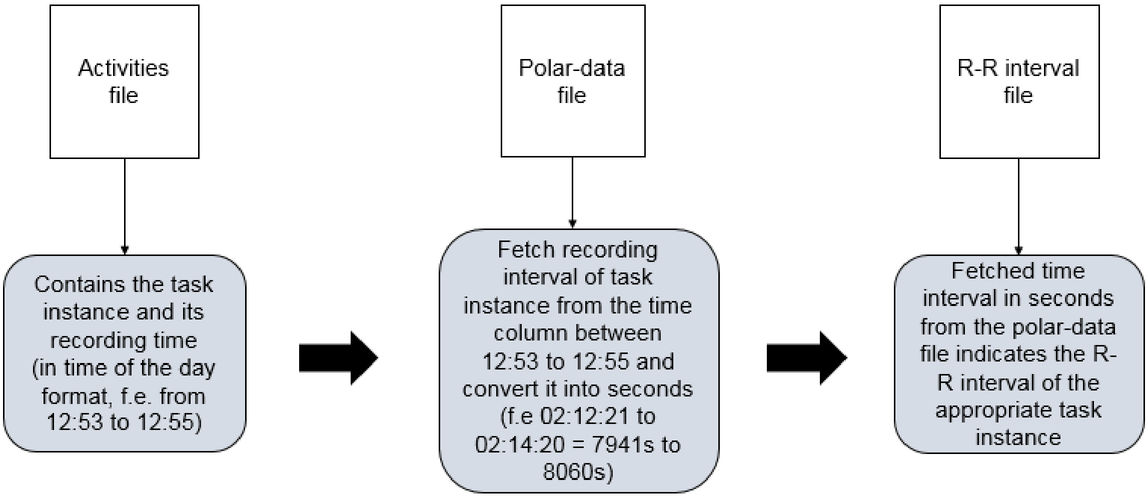


Figure 9: Algorithm for retrieving the R-R interval of a task instance

As it can be seen in the activities file, there exist 181 task instances for the six participants in this study. Therefore, analysing them manually with the above algorithm and for example, Microsoft Excel, would be very time consuming, especially with regard to future larger samples. Nevertheless, the time-domain method was manually applied for Nurse 1, Nurse 2 and Nurse 4 for the task 3 in order to evaluate their heart rate variability and show a proof of concept for the future analysis of the data with the system. In the examples below, also the video material of the participants was analysed to see if reasons for the existing outliers in the graphs could be detected.

Starting with Nurse 1, the heart rate variability for task three is shown below:

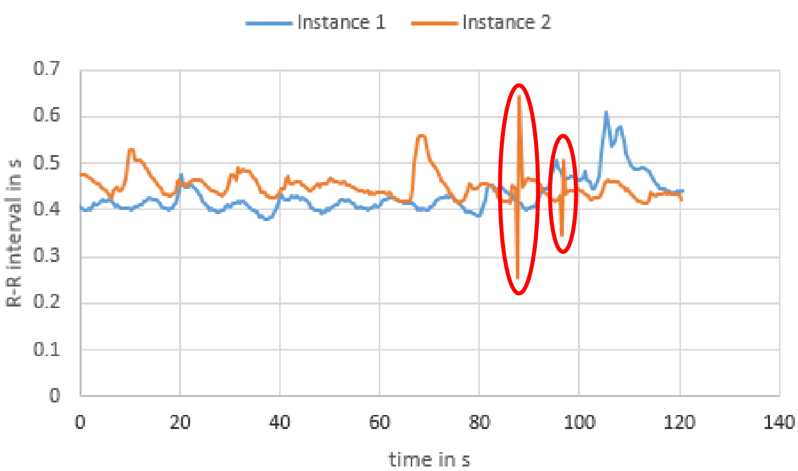


Figure 10: Nurse 1 - Task 3 (with outliers)

Table 3: Nurse 1 - Time domain measures - after cleaning of outliers (∆ >20%)

Instance 1	Instance 2
Mean R-R: 0.428ms	Mean R-R: 0.450ms
SD(RR): 0.036ms	SD(RR): 0.025ms
RMSSD: 0.026ms	RMSSD: 0.030ms
pNN50: 0.3%	pNN50: 0.7%

With a look at the existing video material for the analysis of both outliers, a reason could be a stronger upper body movement and resulting movement artefacts within the chest compression task between minute 14:31 and 14:42, which would match the time in the graph and could thus explain the outliers in the graph.

The following diagram shows the heart rate variability of Nurse 2 for task three:

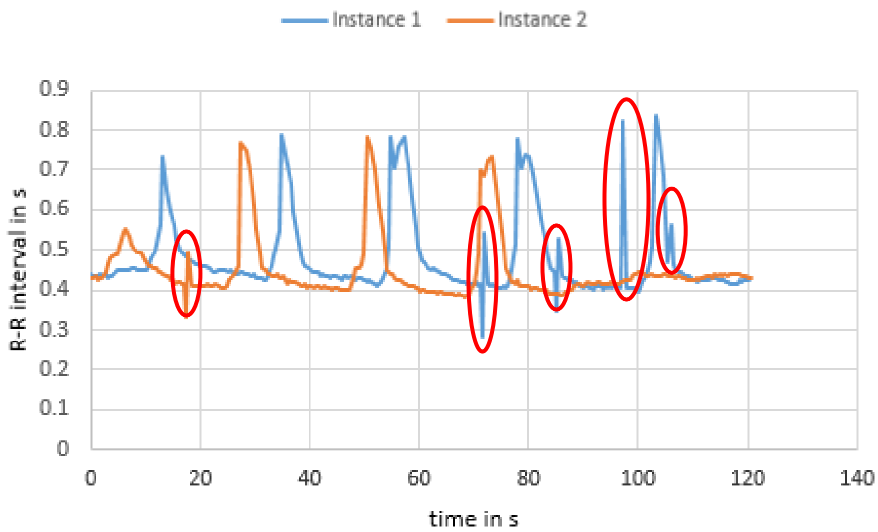


Figure 11: Nurse 2 - Task 3 (with outliers)

Table 4: Nurse 2 - Time domain measures - after cleaning of outliers (∆ >20%)

Instance 1	Instance 2
Mean R-R: 0.458ms	Mean R-R: 0.437ms
SD(RR): 0.072ms	SD(RR): 0.060ms
RMSSD: 0.045ms	RMSSD: 0.041ms
pNN50: 0.7%	pNN50: 0.5%

By looking at the video material of Nurse 2, it could be seen that the nurse gave a chest compression and a mouth-to-mouth resuscitation to the patient in alternate instances, which might lead to the varying peaks and low phases in the R-R graph (mouth-to-mouth resuscitation with deep breathing clearly lowers the heart rate variability

(Sweetwater, 2011)). Nevertheless, it looks like the first three outliers in instance 1 appear through a bending of the upper body as 4:07, 4:24 or 4:31 indicate, right after or before the mouth-to-mouth resuscitation. The outlier in the graph at second 105 in the first instance seems to be a movement artefact caused by an abrupt movement of the nurse while giving the chest compression. In the second instance, the time 10:48 in the video matches second 18 in the graph. In that second, the nurse also bends her upper body over the patient, which might lead to a moving of the sensor and a resulting artefact in the data.

The following diagram shows the heart rate variability of Nurse 4 for task three:

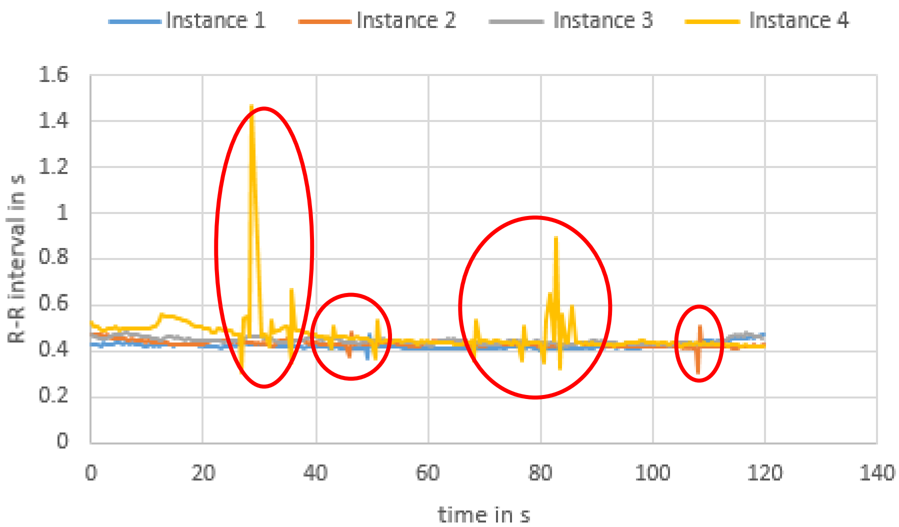


Figure 12: Nurse 4 - Task 3 (with outliers)

Table 5: Nurse 4 - Time domain measures - after cleaning of outliers (∆ >20%)

Instance 1	**Instance 2**	**Instance 3**	**Instance 4**
Mean R-R: 0.423ms	Mean R-R: 0.432ms	Mean R-R: 0.446ms	Mean R-R: 0.456ms
SD(RR): 0.013ms	SD(RR): 0.012ms	SD(RR): 0.013ms	SD(RR): 0.035ms
RMSSD: 0.026ms	RMSSD: 0.029ms	RMSSD: 0.029ms	RMSSD: 0.042ms
pNN50: 0.7%	pNN50: 0%	pNN50: 0%	pNN50: 2.9%

Task instance 4 shows the highest number of outliers among the four instances. The time 27:53 to 28:05 in the chest compression video of cohort 2 represents seconds 26 to 38 of instance 4 in the graph, and also contains abrupt movements during the compression activity which might be an indicator for the outlier values. The time 28:32 to 28:52 in the video roughly represents seconds 65 to 85 in the graph and also shows abrupt pressing movements which might be an indicator for movement artefacts. Nevertheless, during instance 4 in the video, the participant also touches the sensor (see 28:32 in the chest compression video of cohort 2, which also matches the outlier at second 68 in the graph), that might also be an indicator of a loose skin contact with the sensor throughout the whole recording instance and the resulting high amount of error values in the data.

In literature there exist various studies measuring the heart rate variability under different conditions (see section 4.1). However, studies who deal with HRV often refer to difficulties in the interpretation of the results due to the different methodologies used and the setup of the studies (McNarry & Lewis, 2012; Aubert et. al, 2003). Therefore, the obtained time-domain statistics from the examples above are difficult to compare with other studies due to their timeframe (2 min recording) and condition (measurement under physical activity). In literature, most studies conducted which deal with HRV measures normally recommend and conduct recordings within a time frame of at least 5 minutes (Billman, 2011; Dekker et. al, 2000; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Here literature mentions that it is inappropriate to compare measures of various durations (Aubert et al., 2003; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996; Von Borell et al., 2007). Also, with regard to the time-table sheet of this study (see appendix 1.4), it can be seen that some task instances were conducted unequally by the participants with regard to the time duration, which has not always been performed over 2 minutes (as already mentioned that the task instances were sometimes two, sometimes four minutes long). Additionally, the order of the task instances was not the same in every recorded participant, which may also result in inconsistencies within the data. All this lowers the comparability of the results of the different R-R instances and should be avoided in future research.

Nevertheless, the study carried out by McNarry and Lewis (2012) can be taken into closer consideration as an indicator for the aforementioned time domain measures. The study shows, that the more intense the activity gets, the lower the values are for the standard deviation (SDNN) and the RMSSD due to the higher exposure. Although the intensity is always the same for each activity in this study, the individuals might experience different levels of exposure within different instances, which could result in the distinctive values. Apart from that, van Amselvoort et. al, (2000) research concludes that a lower standard deviation of the R-R interval might indicate a less favourable health condition due to the low heart rate variability. Other studies indicate that higher pNN50 values are correlated with stronger parasympathetic activities and greater experienced balance (Dekker, 2000; Taelman et. al, 2009; Vanderlei, Silva, Pastre, Azevedo, & Godoy, 2008; Park, Lee, & Jeon, 2009).

Therefore, with a look at the examples above it can be seen that the instances of task 3 for Nurse 1 show similar values in comparison to each other as well as both instances compared of Nurse 2. Nevertheless, Nurse 4 seems to have a very low heart rate variability for the first three instances compared to the fourth instance of that activity. Also, the statistical measures confirm that the variability is rather low during the first three instances and that the individual might experience stronger levels of exposure. However, as mentioned above; with regard to other studies which recorded data over longer periods such as 5 minutes and more, the results showed higher values and higher significance within the comparison of results among participants (Aubert et. al, 2003; Dekker, 2000; McNarry & Lewis, 2012).

Nevertheless, when comparing the HRV of the three nurses, it is obvious that Nurse 4 has the overall lowest heart rate variability. When looking at the socio-demographic characteristics and health factors of the three individuals, it can be seen that Nurse 4 is the only smoker among them and furthermore has twice or three times a week a standard alcoholic drink. In comparison, Nurse 1 and Nurse 2 – which have more variability in their heart rate, are non- smokers and drink alcohol once per month or less (see digital appendix, participant file). This is an interesting finding which raises questions about the other participants in the study, and how the heart-rate variability is strongly linked to socio-demographic characteristics and health factors.

All in all, it can be said that the recording of the HRV parameters during exercise is insufficiently supported by studies (McNarry & Lewis, 2012). However, a higher HRV which is shown by a higher SD, RMSSD and pnn50 value, is anticipated with better physical resilience against stressors and indicates stronger parasympathetic activities, which lead to a greater experienced balance by the individual (see section 2.1.1; Aubert et. al, 2003; Sweetwater, 2011; Dekker, 2000; Taelman et. al, 2009; van Amselvoort et. al, 2000; Billman, 2011; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996).

4.3 Resulting Requirements

The following section shows an overview of the requirements for the prototype system to be developed in this project; which are determined based on the previous evaluation of the data as well as on project meetings with the researchers conducting the study:

1. The user should have access to a subject view which provides with graphs of the raw R-R interval for a chosen task and its instances for a specific participant, and checks the intra subject reliability of the task. The user should have access to the video showing the participant during that task instance in order to analyse the graph in combination with the activities done by the participant in the video (f.e. looking for reasons for outliers in the graph).
2. The user should have access to a task view which provides with graphs of the raw R-R interval for a chosen task and its instances over all the participants, and checks the inter subject reliability of that task. The user should have access to the video material of each participant as in the subject view above in order to analyse the graph and the appropriate video material in combination.
3. The polar data file (appendix 1.3) of each participant has to be pre-processed. A further column has to be added which relates the task and its appropriate instance to the recorded data row to be quickly able to see which rows adhere to which task instance (see section 5.2.3). This polar data file should be displayed in the subject view.
4. The tool should provide time-domain measures (Mean, SD, RMSSD, pNN50) in the subject as well as the task view for each task instance of the task to be analysed.
5. The user should have access to a validation view where the whole R-R interval of a participant is displayed in order to get an overview of the overall raw R-R interval data. Also a cleaning functionality should be implemented which cleans existing outliers in the data.
6. The user should be able to delete any R-R interval which falls into the range of either:
    - The difference between the R-R interval and its preceding interval is greater than 20%.

- The difference between the R-R interval and its preceding interval is greater than 50%.
- The difference between the R-R interval and its preceding interval is greater than 200ms.

The selected cleaning parameter has to be chosen prior to the analysis and implemented in the cleaning functionality of the validation view.

7. The participants in the videos should be anonymised prior to the analysis of the video material.
8. The user should have access to an overview of the participants' evaluation of the usability of the Polar instrument used in the study.

4.4 Conclusion

Within this work, the time-domain analysis method with its measures of SD, Mean R-R interval, rMSSD and pNN50 was chosen after a comparison with existing literature about the analysis of heart rate variability. With regard to the use of the video material, the user is able to compare seconds of the time-domain graph with the movements done in the video by a participant within those seconds.

With regard to the recording reliability of the Polar H7 instrument, Vanderlei et. al, (2008) and Junqueira and Porto (2009) – who conducted a study comparing another Polar instrument (Polar S810) with the conventional ECG; the polar device was seen as a reliable instrument due to the fact that it provided proper recording of R-R series at rest and exercise over various instances which could be later analysed within the time and frequency domains. As shown in the previous sections, the Polar H7 instrument used in this study can be seen as reliable in the sense that it is measuring data with only minor drop out values over all six participants. Also, a look at the R-R interval recording implies that the instrument records data over all six participants reliably with a maximum outlier rate of 3.33% for participant 1 (see appendix 3.2). Nevertheless, there can be different reasons for the appearance of outliers, as mentioned in the literature review. This makes it difficult to set guidelines which state how many outliers are "normal". The use of video material might help to detect movement artefacts as reasons for these outliers, and is also beneficial to monitor low heart rate variability periods in the R-R interval graph to analyse where the participant experienced higher exposure and stress levels. However, with regard to the existing outliers in the examples in section 4.2.4, movement artefacts (upper body movements and abrupt body movements) seem to be the reasons for outliers during the chest compression activity of the examples shown above.

With regard to the 181 instances of the small population of six participants in this study, the manual collection and plotting of the data as shown above can be seen as too time-consuming and inconvenient. This confirms the requirements for a more time efficient monitoring of the data with the help of a software tool in order to be able to visualise, analyse, and compare the data with regard to outlier values and heart rate variability measures. In addition to this, the different setup of the study (normally 2 min to 4 min intervals under exercise condition) makes it difficult to compare the time-domain measures with other study results using higher time-frames starting from 5 minutes onwards. With regard to this, further research should consider this fact and accordingly adapt the length of the recording time to the setup of another study the results could be compared against.

Nevertheless, with a look to the evaluation of the raw data in this chapter, questions arise which should be analysed with the help of the system in further research as:

- How accurate is the Polar H7 instrument used in this study with regard to the measured R-R interval and the speed and distance measurements?
- What is the reason for the continuous 0 value for acceleration in the data, does a software error exist which is causing these phenomena?
- Is it possible to explain the reason for a high number of outliers in the recorded R-R interval of a participant?
    - In regard to participant 1 and 4, the outlier number was higher than in the recordings of the other 4 participants (see appendix 3.2).
        - Is it because the sensor did not fit properly?
        - Do participants who have a high number of outliers often touched the sensor during the recording time?
- How well does the video material confirm the reasons for outliers in the R-R interval apart from the above shown examples?
- Is it possible to prevent the outliers in the R-R interval appearing through movement artefacts for the next study? (for instance, providing better

instructions to the participants about not to make abrupt upper body movements during the recording period).

- How well does the video material confirm the different experienced exposures and stress levels in the heart rate variability of the participants?
- How do the socio-demographic characteristics and health factors influence the heart rate variability? (comparison of different participant profiles needed)
- Is the data reproducible within an individual over longitudinal studies?

Based on the requirements and questions on this chapter, the following chapter designs a system which should help users to answer these questions and serve as a framework for future research.

5 Design

This chapter outlines the design of the prototype system for the integration, visualisation, cleaning, and analysis of the different data existing in this project. Therefore, two user groups are determined based on the discussions from previous project meetings with the researchers of the study. Furthermore, the first prototype design of the tool is going to be developed with Axure and evaluated by three experts working in this project. The last part shows the software architecture of the evaluated tool to be developed.

5.1 User Identification

Two user groups can be identified in this project, which are on one side, the developers with programming skills; and on the other side, the medical or research staff who is analysing the recorded data. The developers have to be able to deal with Python and maintain the graphical user interface, as well as implement and maintain the different required functionalities of the software tool by writing code scripts. Cooper (2004) suggests to create personas when developing products, especially with regard to a user-friendly design and appropriate coding approach. Therefore, this work includes two persona groups which should be taken into consideration for an efficient development of the software tool:

5.1.1 Researcher

A researcher uses the software to analyse the data to detect patterns and draw conclusions for further research:

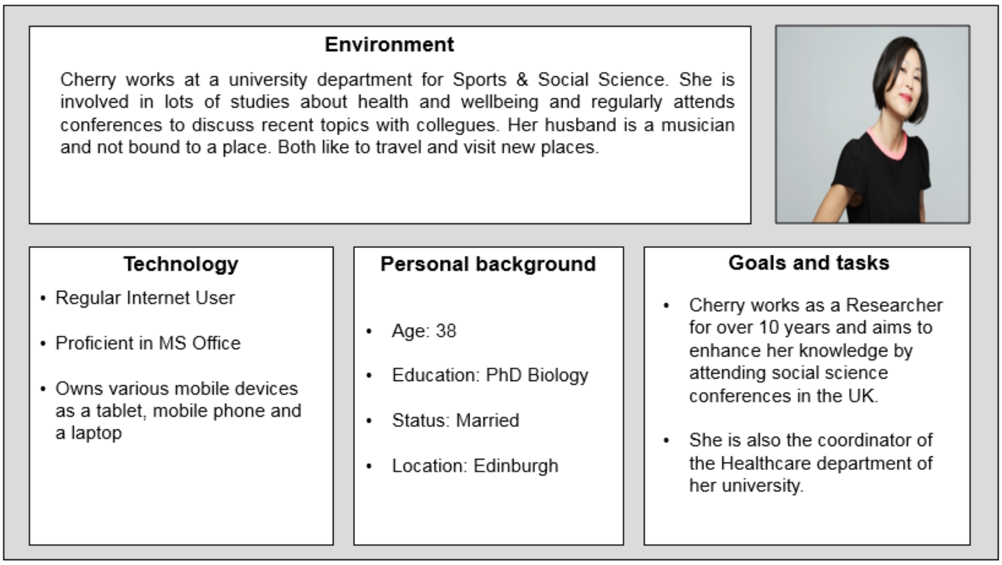


Figure 13: Researcher persona

5.1.2 Developer

A developer develops and maintains the software tool:

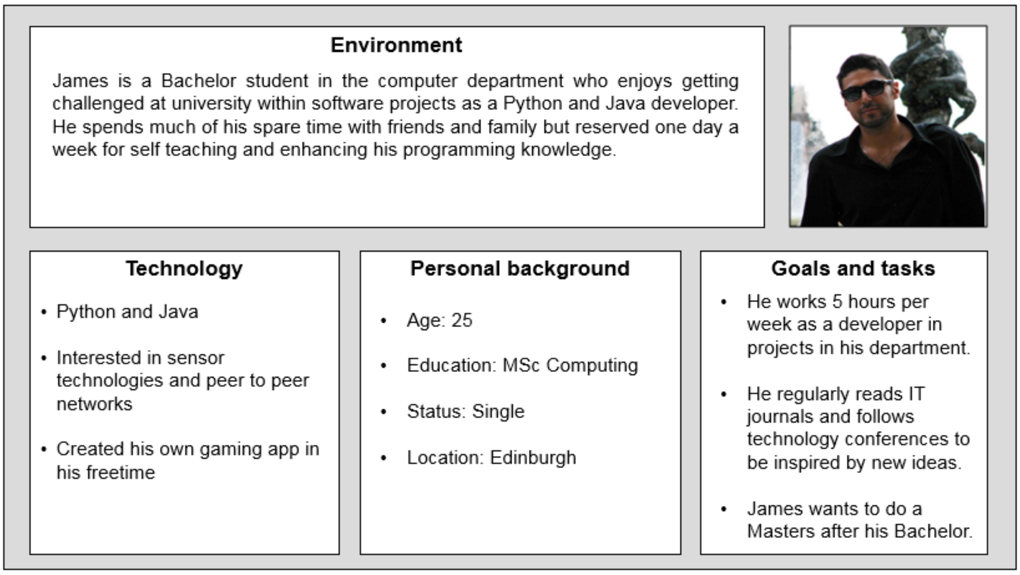


Figure 14: Developer persona

5.2 User Interface Design

With regard to the above mentioned requirements, the following chapter presents the design of a system able to answer questions about the recorded data. The graphical user interface is designed in Axure, an application for fast prototyping which is used to demonstrate the user experience of the software application of this work (Daliot, 2013):

5.2.1 Device Usability

The first screen the user gets to see when accessing the system is a summary of the answers given by the participants about the usability of the instrument utilised in the study; which are retrieved from the participant data file (see appendix 1.1). The answers to the questions and statements show the average over all participants in the study.

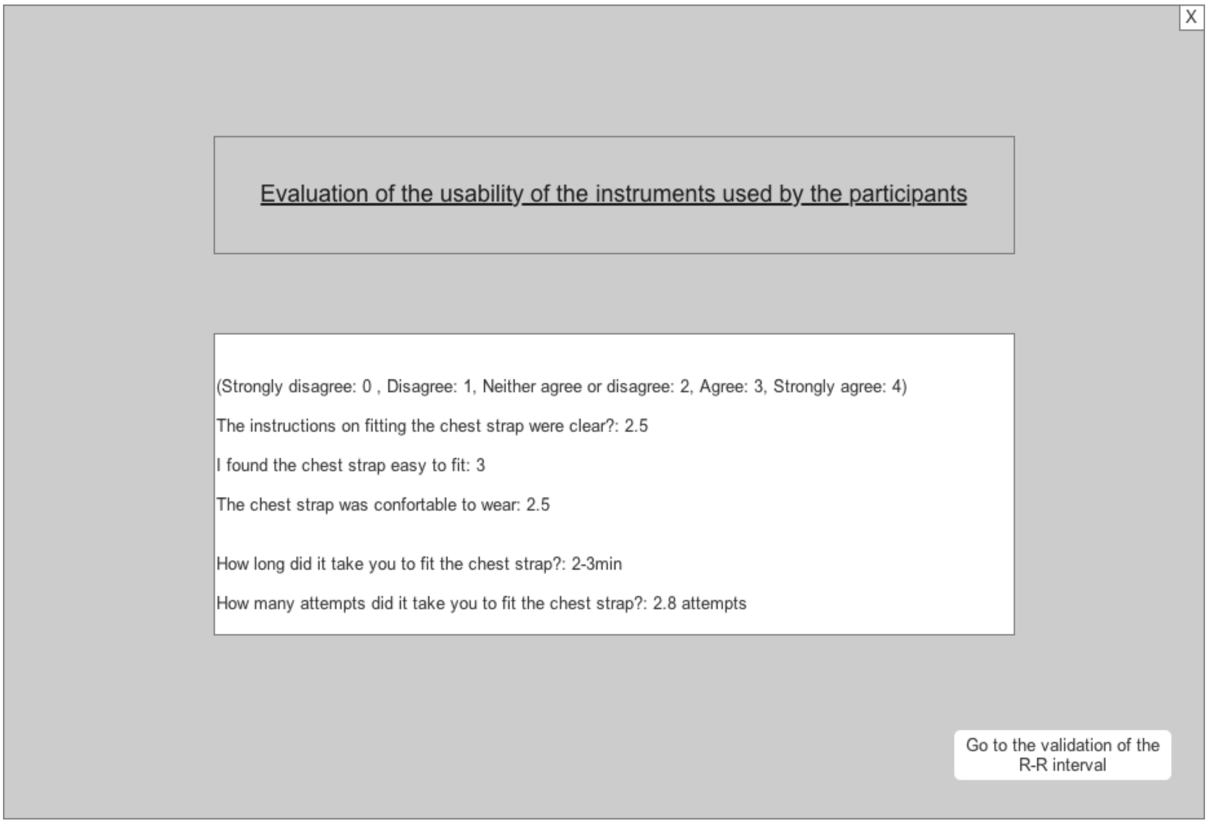


Figure 15: Usability view

The button on the right side of the GUI leads the user to the validation view where the R-R interval of a participant is displayed over the whole recording period (Figure 16).

5.2.2 R-R Interval Validation

The screen containing the validation of the R-R interval includes a button list in the right side which contains the participants of the study. With the selection of a participant, on the top of the screen the socio-demographic characteristics and health factors of the participant (from the participant data file) are shown; and the R-R interval over the whole recording period is displayed. At this stage, the user can zoom into the graph and specifically monitor appearing outliers or other inconsistencies in the R-R interval.

A button on the bottom of the graph would delete outliers in the data which are either greater than 200ms, greater than 50% or greater than 20%. The selection of the additional removal criteria has to be chosen prior to the analysis of the data, and has to be implemented by the developer. Through the deletion of the outliers, the new cleaned graph is displayed. A save button is also implemented which enables the user to store the graphs as a file. Apart from that, on the lower right side of the window, the user can choose to go directly to the task or subject view of the application.

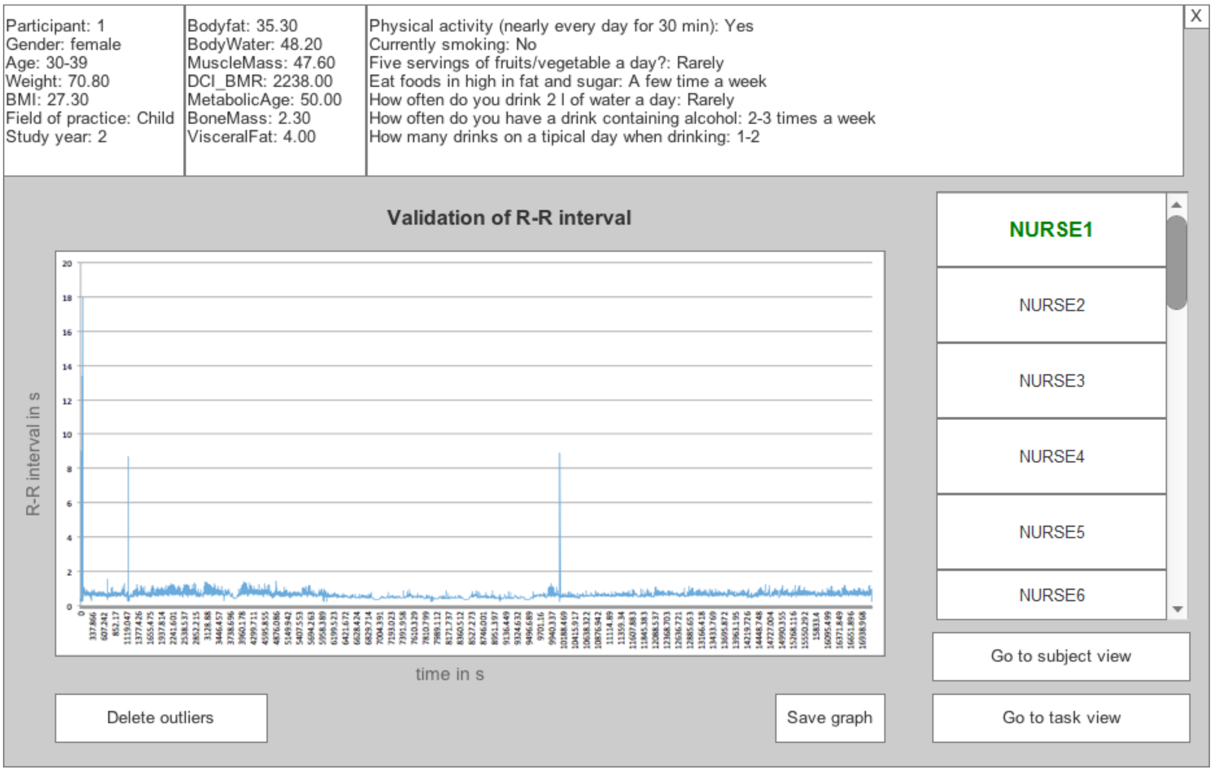


Figure 16: Validation View

5.2.3 Intra Subject Reliability

After clicking on the “Go to subject view” button in Figure 16, the user gets to the subject view. In this view, the user can check the intra subject reliability of the R-R intervals of a task. The subject view is designed to enable the user to select a participant and a task for which the raw R-R graphs should be visualised and analysed. The button list on the right side contains all participants of the study, whereas the 9 buttons on the top of the application window represent the tasks. In addition to that, the graph should be interactive; which means that if video material exists for both task instances, it can then be aligned with the graph (both should have the same recording time so it can be aligned over the seconds). This means the user clicks on a segment in the graph where for example an ectopic interval exists and the video of this segment appears, showing the participant's action in that time period (see section 6.6.2). The user can also press the button "PD", which opens the polar data file containing the heart rate and speed/distance measurements of the participant. The first column "task" in the polar data file shows the task and instance (f.e. Task 3/1) to make it easy for the user to find the data in the file related to the displayed R-R graphs. Furthermore, the time-domain statistics are also shown in a separate window for the selected task in order to evaluate the HRV of the participant. Finally, the graphs can be saved by clicking on the "save the graph" button on the lower right side. The button "Go to task view" directs the user to the task view of the application.

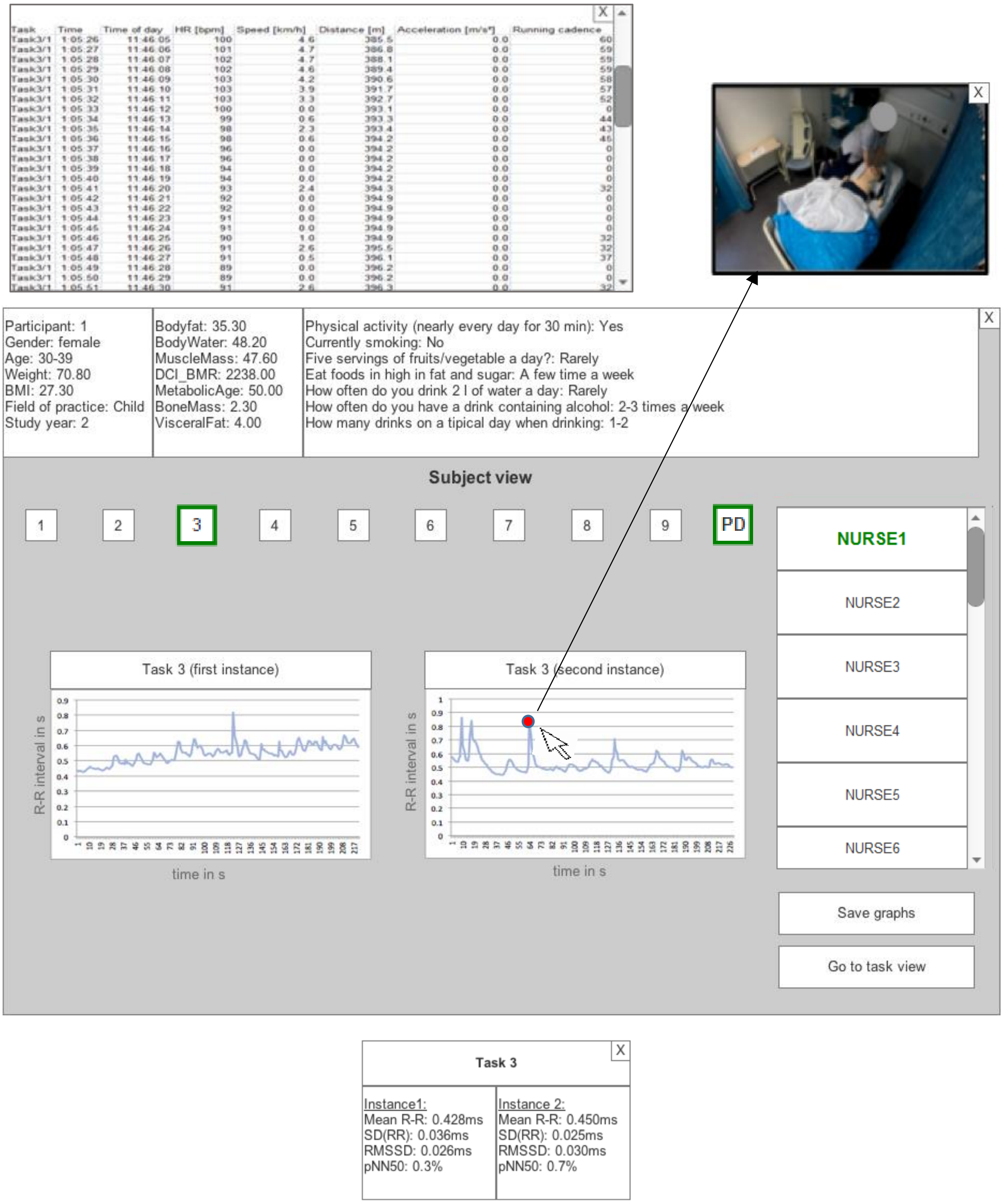


Figure 17: Subject view

5.2.4 Inter Subject Reliability

The task view enables the user to compare the inter subject reliability of the R-R interval over all participants for a selected task. The right side of the application window displays a button list from which the user can choose a task to be analysed with its instances. In this view the user can also click on a graph segment (f.e. an outlier) which opens the video part showing the participant’s activity in that time segment in order to draw conclusions from the movements. Besides, the time-domain measures are calculated over all the participants for the selected task, which assess the heart rate variability of the participants.

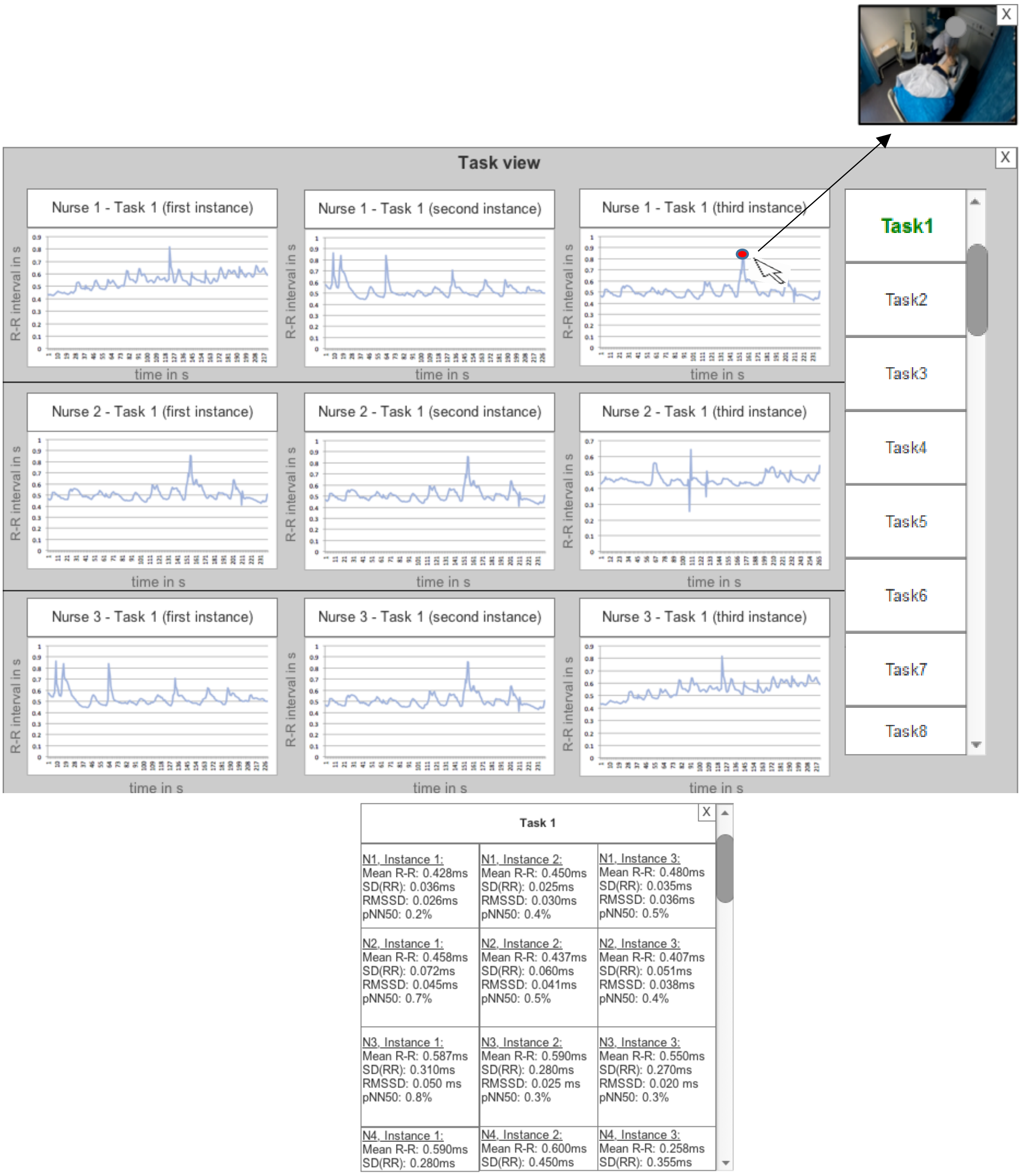


Figure 18: Task view

5.3 User Interface Evaluation

As already mentioned in section 3.4, the developed prototype tool was evaluated by three project experts in the field of healthcare and nursing through a semi-structured interview. The questions of the interview are developed based on Davis (1989) and his

Technology Acceptance Model and with regard to the research questions this work is trying to answer. A brief explanation was given to each interviewee about the different views of the tool with all their functionalities to gain a better understanding and increase trust within the participants (Saunders et. al, 2009). The answers of the interviews revealed significant insights for the development of a robust and beneficial software tool (see appendix 4.1 for the analysed interview and appendix 4.2 for the detailed answers):

With regard to the ease of use and the usability of the tool, a pattern could be detected in the answers: the different steps for each view were commonly described as being very intuitive and easy to follow. The interviewees further described the tool as highly usable with regard to the integrity of all the information displayed in one screen; as well as the possibility to compare the graph of the R-R interval with the activities done in the video by the participant. With regard to the possibility to draw conclusions from the data, all interviewees agreed there is a high potential to identify shift patterns, exhaustion levels or outliers; and that judgments can be made about the heart rate variability based on individual comparisons between the participants. However, as mentioned in one interview, “there are not enough participants” (see appendix 4.2) to make judgements or drawing conclusions from the population in this study.

A task as well as a subject view is seen as highly beneficial by the interviewees due to the fact that the HRV during a task can be compared within one individual (subject view) as well as over different many individuals (task view). Nevertheless, it was mentioned that an overlay of the graphs in the tool would better enable the users to make a more accurate comparison and should then be implemented.

In terms of how the designed tool enables the user to judge the accuracy of the collection equipment, all three interviewees mentioned that there exists no comparison of the resulting data with for instance, another instrument; which makes it difficult to assess the accuracy of the Polar H7 instrument. With regard to the reliability of the Polar H7, a pattern in the interviewees’ answers can be detected, as the majority perceived the instrument as reliable due to the visualisation of the results in the software, that makes the data comparable over the nurses. Furthermore, the tool was seen as reliable because “it works as expected and there are multiple datasets of the same task for a comparison” (see appendix 4.2), statement which matches the findings about the reliability discussed in chapter 4.4. Furthermore, all interviewees agreed that a judgement can be made about the perceived usability by the participants with the help of the software tool.

With regard to which functions might be added to the developed prototype system, the different interviews revealed that an overlay of the graphs in the different views is the key demand of this evaluation to be able to accurately compare the HRV over the various instances and among individuals. Also in the Validation View of the tool (see section 5.2.2, Figure 16), an overlay of the whole R-R intervals was demanded to better be able to compare all the different recordings. In addition, it was mentioned that users of the tool should be able to align the socio-demographic characteristics and health factors of the individuals with the analysis of the heart rate variability data; "it would be beneficial to tag participants based on key socio-demographic characteristic for example in the task view to tic boxes for smoker, fitness activity (.) and nurses would be tagged when there is a fit" (see appendix 4.2). Furthermore, an interviewee mentioned to also plot the time-domain measures of the R-R interval apart from the graphs of the "normal" R-R interval for a task; in order to have an additional option to compare the heart rate variability among the participants.

Another point mentioned was that the user should be able to clean the raw data from outliers not only in the validation view, but also in the task and subject view. Here the evaluation showed that more cleaning possibilities of outliers should be provided. Also, the display of the number of outliers cleaned for an individual was demanded.
With regard to the question about the knowledge of Python, no one of the interviewees did know about the language; what justifies the need of a second user group: developers who are able to quickly generate new Python scripts and maintain the tool.
Another result of the interviews was a discussion about the implementation of frequency domain measures of the HRV in future enhancements of the tool, in order to get the ratio between low frequency over high-frequency as another indicator for the experienced stress level of an individual.

With a look at the questionnaire evaluation (see appendix 4.3), it confirmed the answers given in the interviews by the experts, and showed a high repeatability in their answers (Creswell, 2013; Saunders et. al, 2009). Only two questions revealed inconsistencies. The rating of which functionalities could be added to the existing tool showed, with regard to the interview answers; too positive ratings by two of the interviewees, whereas the rating given by one interviewee for the accuracy of the instrument could be seen as lower than the statement given in the interview before.

These inconsistencies between the questionnaires and interviews might be given due to bias in the interviewer with regard to the explanations provided during the interviews (question about the accuracy), as well as his presence in the same room when the interviewees completed their questionnaires (Saunders et. al, 2009; Creswell, 2013).

Taking into account the above demanded adjustments, the prototype was improved:

5.3.1 Device Usability

The screen containing the participant evaluations of the instrument used in the study was not modified due to no demanded changes (see section 5.2.1).

5.3.2 R-R Interval Validation

In the new screen containing the validation of the R-R interval, buttons with all different cleaning measures (see section 4.2.3) were implemented with regard to the demanded greater flexibility for removing outliers. Now, the user can choose between a more aggressive or less aggressive cleaning method for the displayed raw data (see section 2.1.2). Also the number of outliers each cleaning parameter identifies and cleans in the raw data after pressing the button is displayed under each of the parameter buttons. This enables the user to analyse the raw R-R graph and the cleaned R-R graph of a participant (cleaned graph is displayed after choosing a cleaning measure). Furthermore, a button is implemented on the bottom side of the window, including a function which is opening a side window displaying all R-R intervals of all participants in one diagram, in order to have an additional option to compare the recorded R-R graphs among participants.

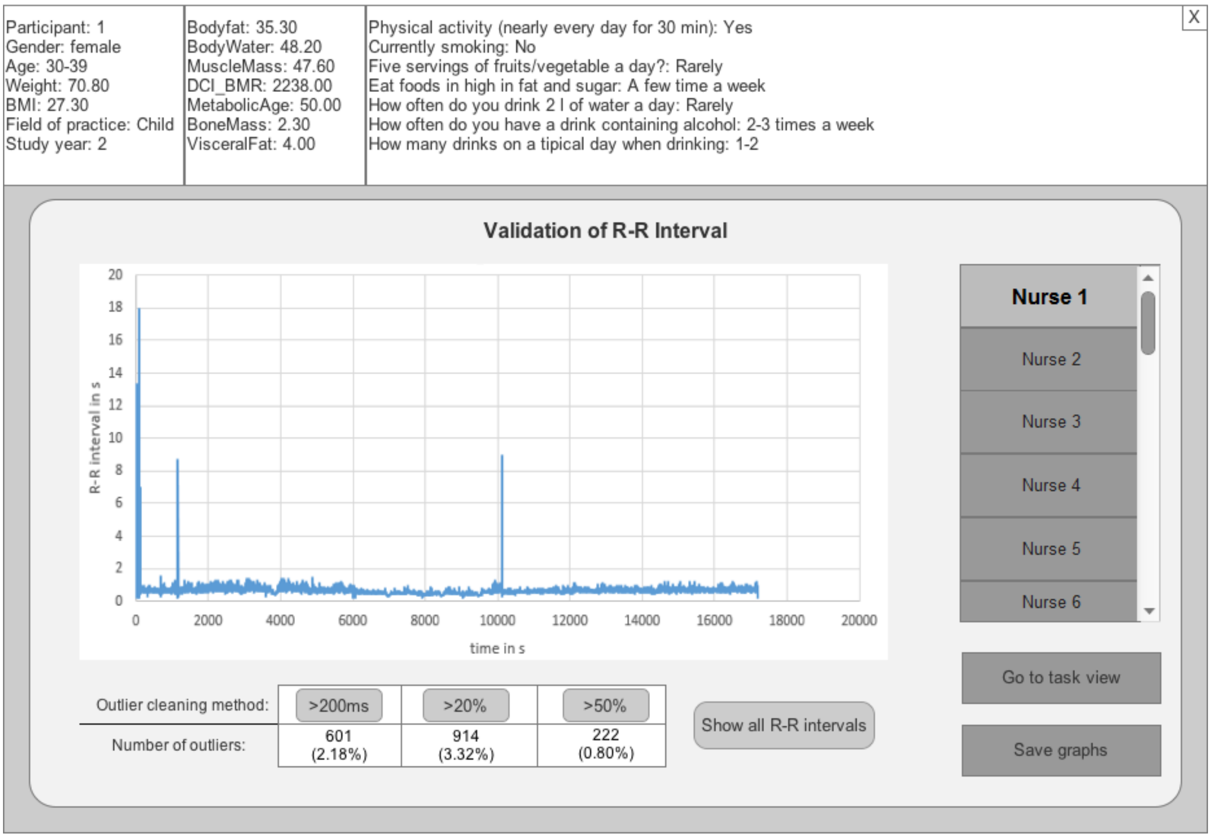


Figure 19: Evaluated Validation View

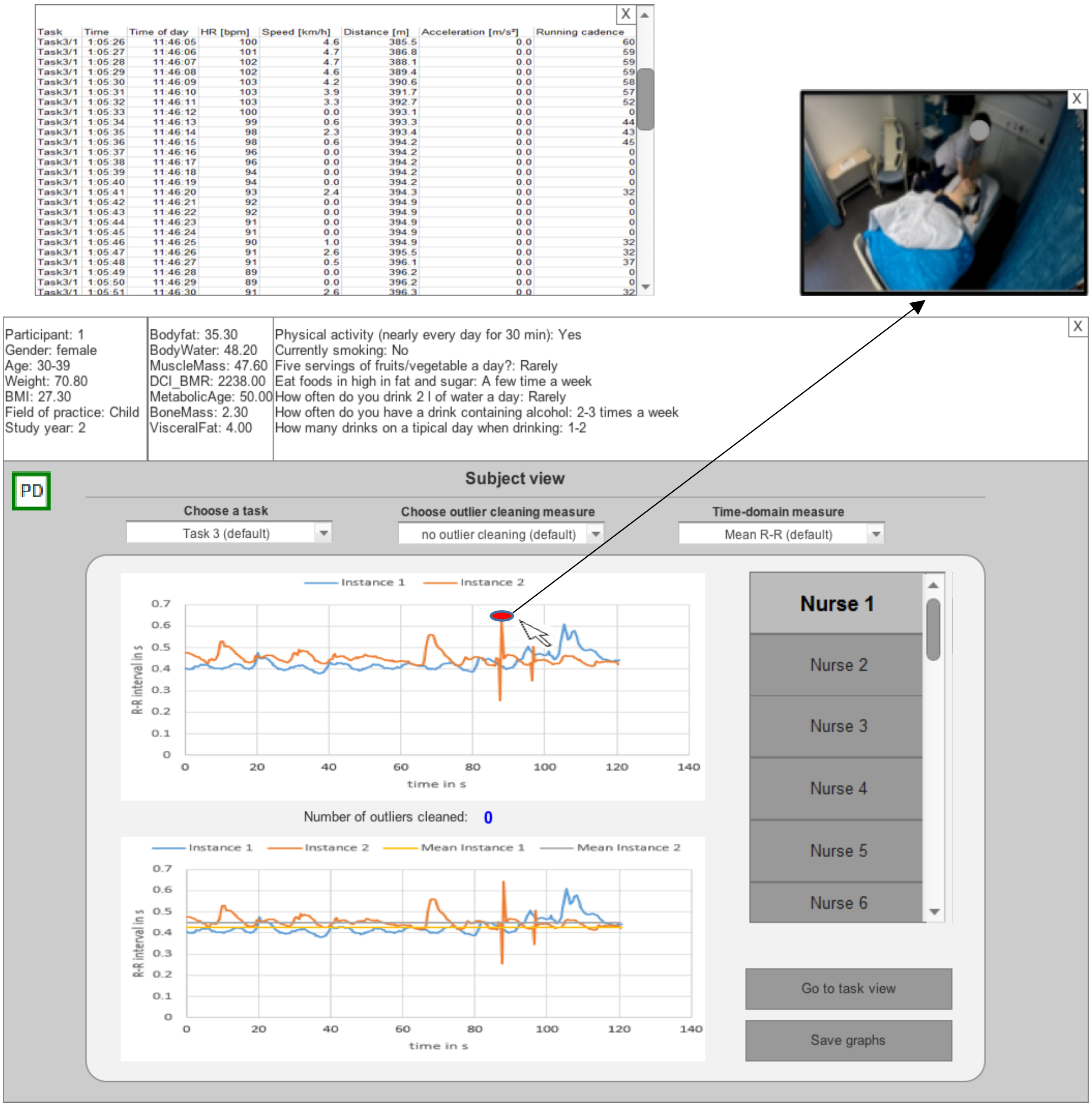
5.3.3 Intra Subject Reliability

Modifications were also made in the subject view which compares the intra subject reliability of a task. To the subject view three drop-down lists were added, one for the selection of the task which should be analysed, one for the cleaning measure which should be applied, and one more for the time-domain measure of the selected task that should be plotted. Furthermore, all R-R interval instances are now displayed within one diagram window, making the heart rate variability of the R-R interval better comparable among the task instances. In addition to that, the polar data ("PD" button) can be displayed to compare the graph with the other measures of the instrument as shown in the first prototype screen (see section 5.2.3, Figure 17). Below the graph of the R-R interval instances for the selected task, there is a new diagram window showing the chosen time-domain measures for the R-R interval instances of the selected task, as demanded in the interviews conducted with experts. The user can choose within the drop-down lists:

- one task from the nine tasks of the study

- an outlier cleaning measure (no cleaning, >200ms, >20%, >50%)
- a time-domain measure (Mean, SD, RMSSD, pNN50) to be displayed for the selected task.

By selecting a value from the drop-down list, the graphs are automatically adapted to the selected value. The number of outliers cleaned is displayed in the middle between the two diagrams. As in the screen of the first prototype design, the R-R graph is interactive and a click on a segment displays the video showing the participant activities for that moment (see section 6.6.2). Furthermore, the time-domain statistics of the instances of the chosen task are automatically displayed in a side window below the GUI.



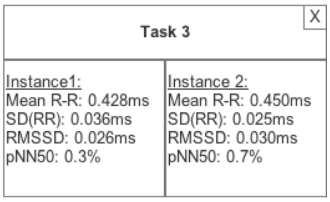


Figure 20: Evaluated Subject View

5.3.4 Inter Subject Reliability

The adjusted task view represents the inter subject reliability of a task and contains as the subject view the R-R interval (left graph) and the time-domain measure of the R-R interval (right graph) of a specific task. The difference to the subject view is that it displays the data over all the participants of the study.

In this view, also three drop-down lists were added, one for the selection of a task which should be analysed, one for the cleaning measure which should be applied, and one more for the time-domain measure of the selected task that should be displayed. The selection options are the same as explained above in the subject view. These are applied for all the study participants in order to be able to compare the heart rate variability of a chosen task over all the participants. Therefore, by selecting a value from the drop-down list, the graphs of all the participants are automatically adapted to the value selected. As in the subject view before, the R-R interval graph can be aligned to the video material of the task instance as well. Also, all the time domain-statistics are displayed in a window below for the selected task and its instances over all the participants of the study.

Furthermore, socio-demographic factors such as smoking, physical activity or age (chosen as examples from the participant data file) are integrated in the analysis process. The user can tick one or more criteria to investigate (here for example “physical active”) the socio-demographic characteristics, which will mark the participants who match the criteria in red (in the example participant 1 is physically active and tagged red). Apart from that, the number of cleaned outliers for the data is displayed for each participant as well.

Figure 21: Evaluated Task View

5.4 Software Architecture

The software architecture demonstrates the foundation of the operational functionality of the tool. Therefore, it shows the proposed distribution of the recorded data and its appropriate script files which include the code for the querying of the data, the integration, the cleaning and visualisation of them. First, the folder structure of the tool is shown, outlining the setup of its folders which contain the raw data and the required code scripts. Apart from that, the proposed software architecture is explained followed by a proof of concept of its simple maintainability and robustness.

5.4.1 Folder structure

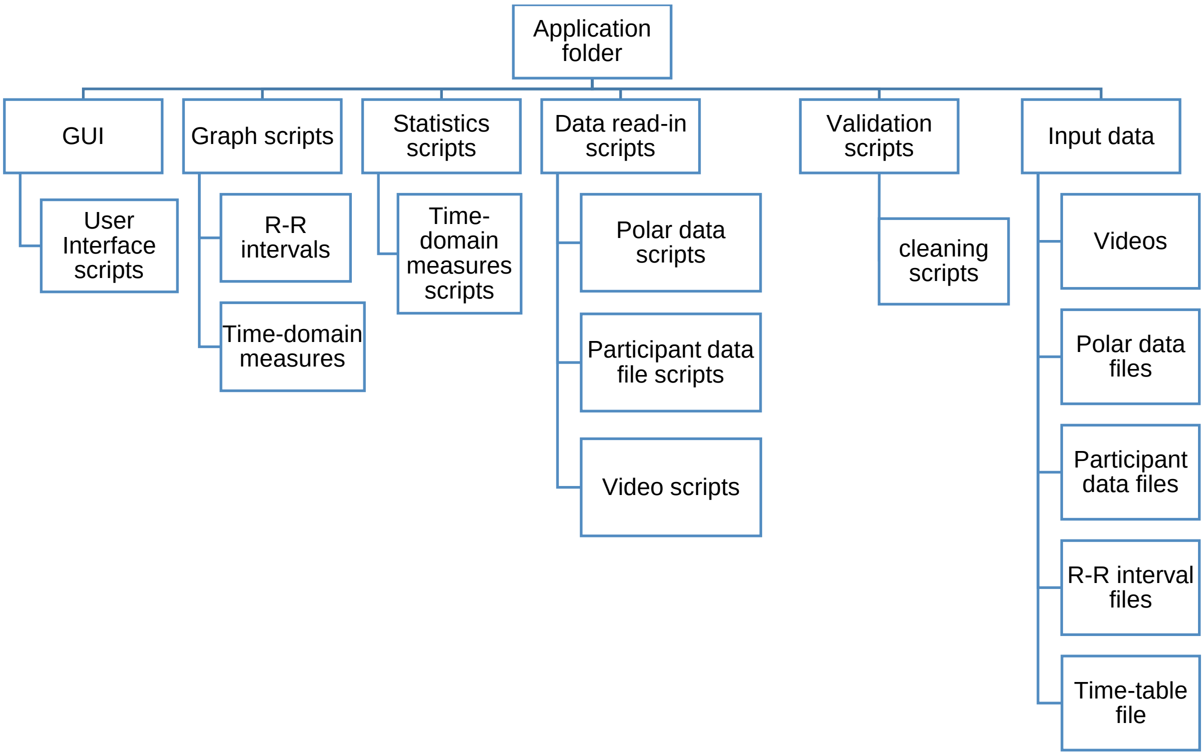


Figure 22: Folder structure of the system

The folder structure of the application shows the distribution of the scripts and data files to be used by the prototype system. The overall application folder consists of six sub folders:

the GUI folder consists of scripts which shape the GUI such as buttons, tic-boxes, drop down lists, or the allocation of the data to fields in the interface. Apart, the graph scripts folder contains all scripts which are responsible for the visualisation of the retrieved R-R intervals from the participants' R-R interval files (see section 5.3, the R-R interval is plotted as well as the graph of its time-domain measure). Apart from that, a statistics folder exists, which contains the scripts in order to measure the time-domain statistics for a task and its instances from the participants' R-R interval. The data read-in folder consists of the scripts which read the answers about the usability evaluation given by the participants (see section 5.2.1) as well as the socio-demographic characteristics from the participant data file. Furthermore, the Polar data with the heart rate and time measures for a specific participant is read from the polar data file. In addition to this, there exists a script which displays the appropriate part of the video with regard to the graph segment the participant clicks on. The fifth folder in the structure is a validation script folder, containing scripts with the different options for the outlier cleaning of the R-R interval. Last but not least, the input data folder is used to store the raw data for each participant as well as the time-table file (see appendix 1.4).

5.4.2 Proposed Software Architecture

In the following chapter, the overall concept of the software architecture proposed for the prototype to be developed in this work will be shown. From the folder structure (Figure 22), the software architecture can be derived, showing the different data elements of the application and where they are retrieved from. As already mentioned in the literature review, a big advantage of Python is that its scripts are not compiled and can therefore be easily changed or extended. This means that a developer can write scripts with a simple text editor and store them in the appropriate folder. The architecture's folders contain various scripts in order to perform the different required functionalities retrieved from the previous chapters.

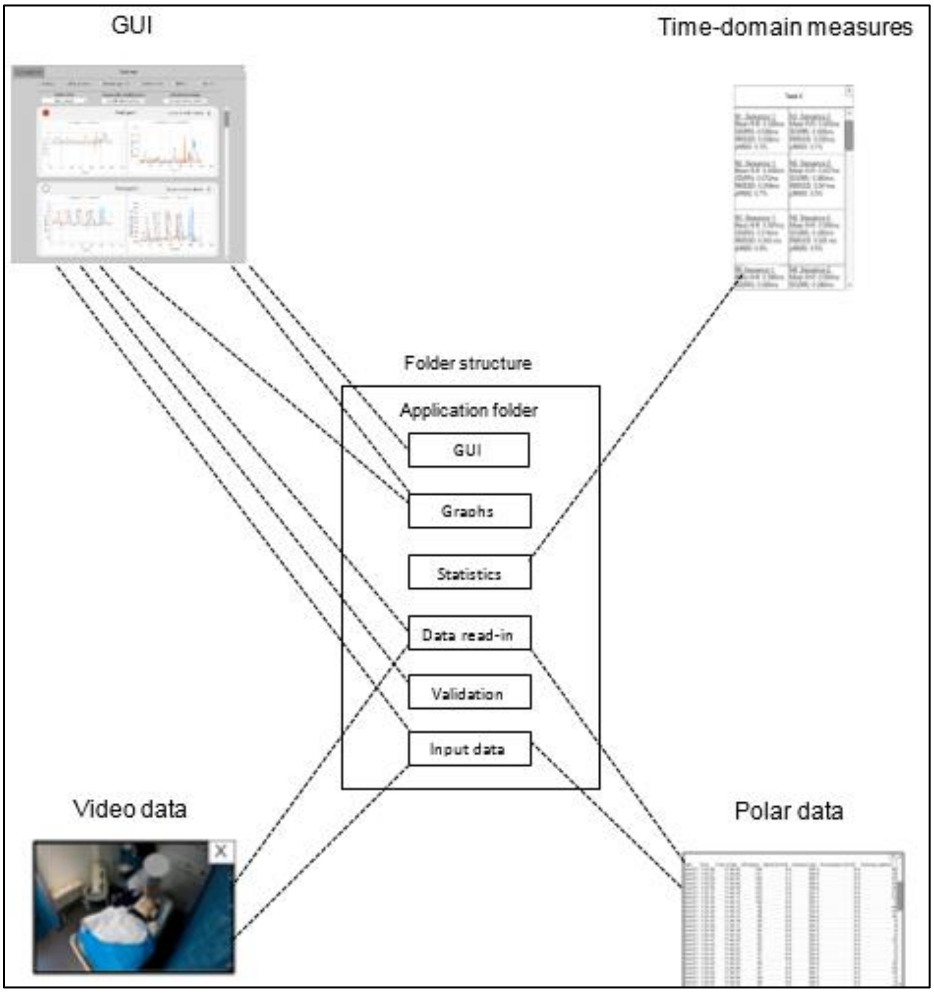


Figure 23: Proposed Software architecture

5.4.3 Software Architecture Robustness and Maintainability

The simple maintainability and robustness of the system against total breakdowns caused by script errors is shown through a comparison between two examples. The first example (see Figure 24), shows a software architecture including a folder structure containing only one script file that stores all code. In this case, there are less files and a simpler folder structure than the one displayed in Figure 22. Nevertheless, overall the possibility for total breakdowns is higher than with a more distributed solution. A single error in the script leads to a crash of the GUI and other parts (video, polar data, time-domain measures) created by the script.

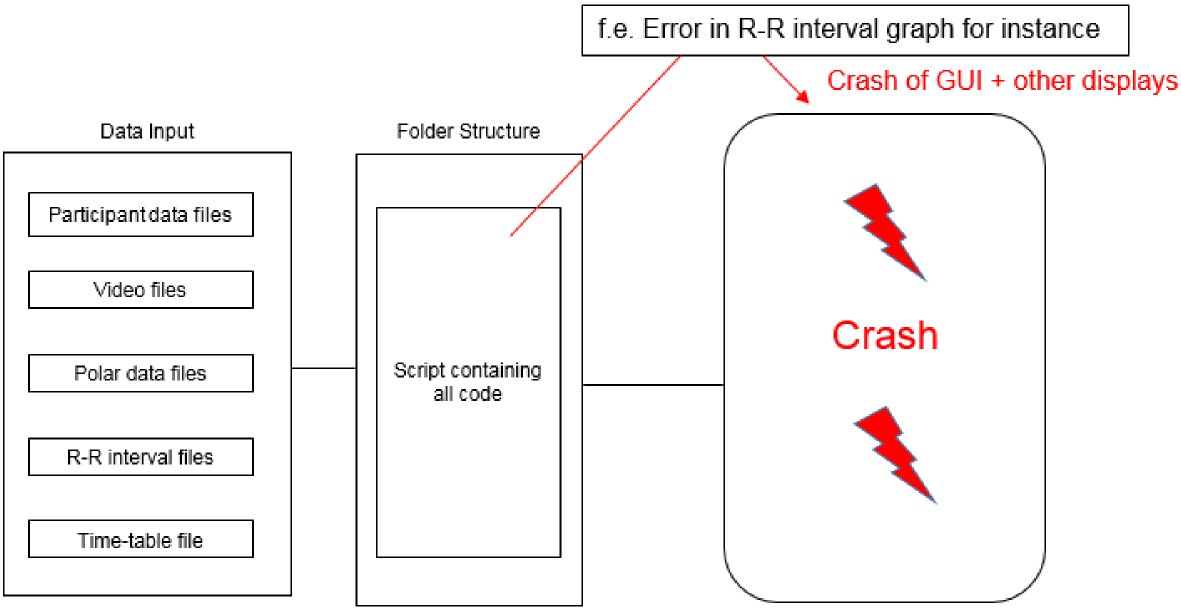


Figure 24: Software architecture with only one script file

Conversely, Figure 25 shows the advantage of the already proposed architecture for the system, which includes code distributed over several scripts so each script can be easily manipulated without leading to a total crash of the whole program. If an error is appearing in a script this will only lead to a crash of the function related to the script.

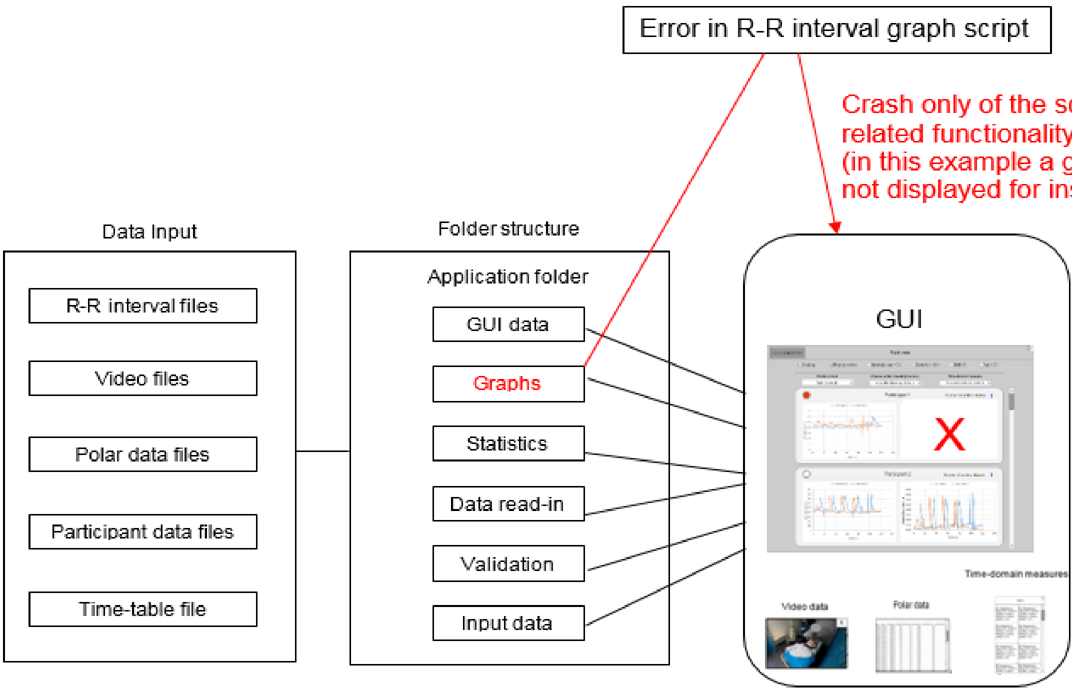


Figure 25: Distributed model of a software architecture

5.5 Conclusion

The system to be developed has two user groups. Developers who maintain the software and researchers who analyse the data. A prototype of the system was designed with regard to the stated requirements of section 4.3 and evaluated with three experts in the field of healthcare. The conducted interviews showed that further requirements had to be implemented (added to section 4.3), as:

- an overlay of the different R-R instances (to be added to requirement 1, 2).
- an integration and alignment of socio-demographic characteristics and heath factors within the analysis of the HRV in the task view (to be added to requirement 2).
- another plot showing the time-domain measures of the R-R interval (to be added to requirement 1 and 2).
- an outlier cleaning possibility in each screen (not only in the validation view of the R-R interval) and more cleaning flexibility. This means the users should be able to choose if they want to clean more aggressively or moderately (to be added to requirement 1 and 2).
- the display of the number of outliers cleaned (to be added to requirement 1, 2 and 5)
- a function in the validation view of the system, which plots the R-R intervals of each participant all together, in order to be able to compare the recordings of all the participants (to be added to requirement 5).

Furthermore, with regard to the research questions of this work, it was revealed that the data recording of the Polar H7 instrument was seen as reliable due to the fact that it is continuously measuring data over all participants, as discussed in Chapter 4. Also referring to what extend the tool might benefit to further research (see appendix 4.1), the designed tool could help to plan and allocate shifts and tasks to nurses more efficiently with regard to their measured stress levels in order to offer the highest care potential. In addition to that, the tool could be used to monitor stress and exhaustion levels in real time, specifically in occupations with high responsibility (nursing, air traffic control etc.). Last but not least, the architecture of the software includes various folders and scripts so each script can be easily manipulated without leading to a total crash of the whole program if it includes errors.

6 Implementation and Testing

Different requirements mentioned in the previous chapters were implemented in Python and tested through acceptance tests in order to evaluate their operational performance. The implemented parts where displayed in the Integrated Development Environment (IDLE) of Python to check their functionality. The focus of this chapter was not set on developing the graphical user interface – due to the limited time frame of this project. Also, the mentioned algorithm for the fetching of data out of the file could not be implemented due to temporal limitations. However, the focus of this chapter is rather set on providing a proof of concept of the key operational functionalities of the system for the analysis of the recorded data. During the writing of the code, several packages and modules were used, as explained in the following sections.
The scripts can be found in the digital appendix of this work (CD-ROM).

6.1 Setup

Python 2.7 was installed for this project as well as “pip” – the Python package which allows the download of python libraries directly with the use of the command line (see Figure 26).

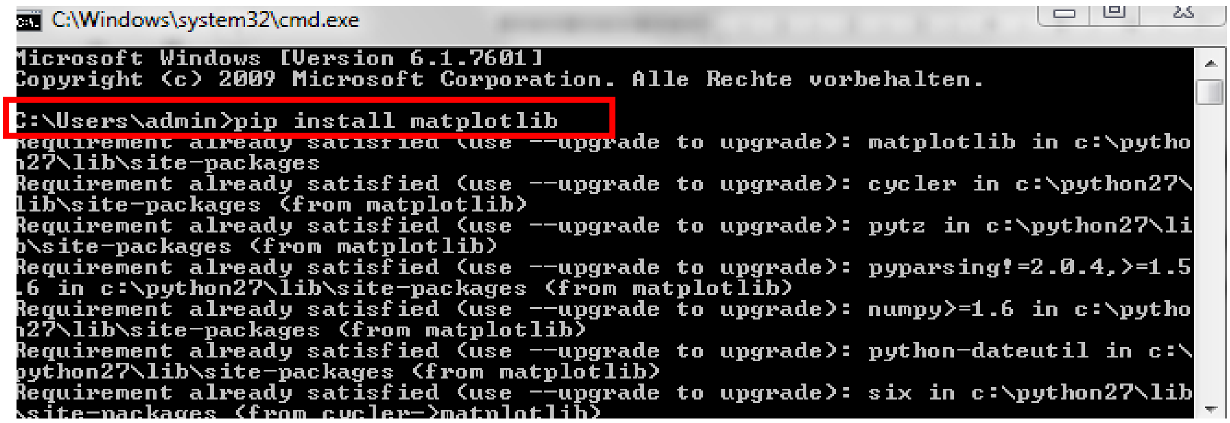


Figure 26: Pip install of packages

Furthermore, all data spreadsheets of the study had to be converted in csv files to read them into Python. The R-R interval files (see appendix 1.2) had to be pre-processed as well in order to allocate headings to their columns, which are necessary for the later processing with Python.

As mentioned above, the results of the scripts were displayed in the embedded Python console and can be opened and executed as follows:

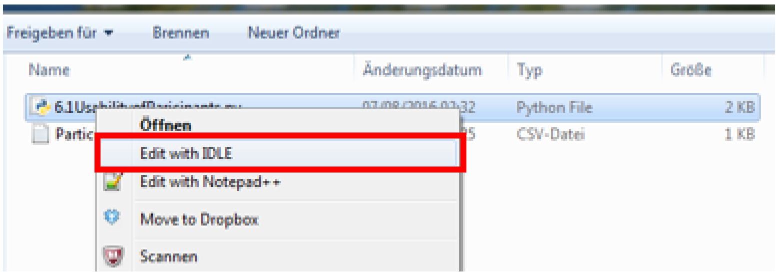


Figure 27: Open script in IDLE

After opening the script file in IDLE, it can then be executed:

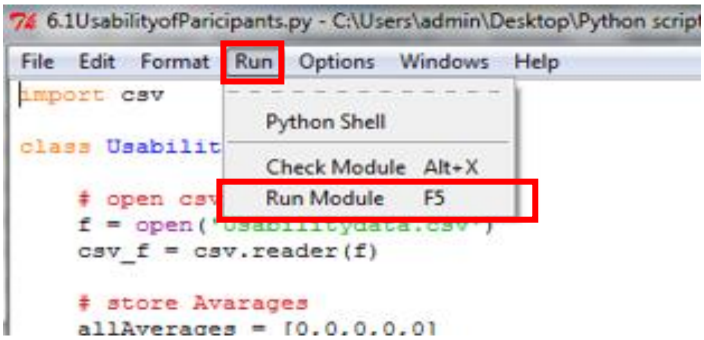


Figure 28: Run script in IDLE

The console is a very helpful tool especially for debugging and testing whether the code is generating the expected results.

The following sections therefore show a proof of concept for the implementation of requirements 1, 2, 3, 5, 6 and 8 of section 4.3 and a discussion of the requirements 4 and 7.

6.2 R-R Interval Visualisation

Chapter 4 discussed the requirements of the software tool to be developed in Python. With regard to the requirements 1, 2 and 5 (see section 4.3), the visualisation of the R-R interval over a task instance was demanded. The following shows the proof of concept for plotting the recorded data in Python. First, the csv module of Python had to be imported in order to read the file. In addition to this, the library matplotlib with pyplot was imported to plot the data of the file.

Therefore, an array was created which saved the created list of objects including the relative time and value of the R-R interval to be plotted in a diagram.

Sample data was created with a smaller amount of records for the testing of the functionality of the script, and to compare the manual approach in MS Excel with the generated plot in Python.

The expected result was:

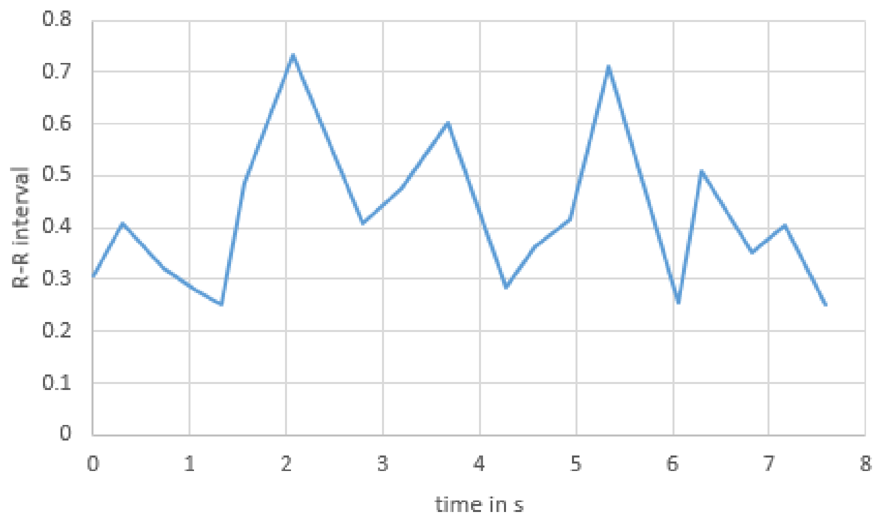


Figure 29: Visualisation of R-R Interval in excel

Result of testing the code: The executed Python code showed the expected output of Figure 29 and the requirement in regard to plotting the R-R interval is satisfied.

As it can be seen from the script, the plotting of graphs in python can be done with just a few lines of code and can easily be integrated in a graphical user interface in a later step.

6.3 R-R Interval Cleaning

As mentioned in the requirement 6. in section 4.3, the user should be able to clean the recorded raw R-R interval data from outliers. Therefore, three cleaning parameters were mentioned in literature (∆ >200ms, ∆ >20%, ∆ >50%) which identified R-R intervals as errors. To implement this requirement, the csv module was imported again to be able to read and open the R-R data file of a participant.

The core function of the script is the iteration through the created arrays if the interval difference between an interval and its preceding interval is within the allowed range:

```
def checkValidInterval_inPercentage (time_1, time_2, percentage):
    if (time_2 / time_1 < 1 - percentage): return False
    if (time_2 / time_1> 1 + percentage) : return False
    else: return True
```

Figure 30: Outlier checking >20% and >50%

```
def checkValidInterval_inMStimespan (time_1, time_2, timespan):
    if (time_2 - time_1 > timespan): return False
    if (time_2 - time_1 < -timespan): return False
    else: return True
```

Figure 31: Outlier checking >200ms

Last but not least, a function was implemented in order to print the detected valid and invalid intervals to have an overview about the result.
An acceptance test was conducted in order to test the operational functionality of the code and check whether it fulfils the requirement 6, stated in section 4.3. Therefore, the R-R interval of Nurse 1 was checked for the first 1000 intervals with the three cleaning parameters. The expected results were **59** error intervals for parameter ">200ms", **88** error intervals for parameter ">20%" and **48** error intervals for parameter ">50%".

Result of testing the three scripts: The executed code shows the expected number of outliers for each parameter and differentiates them from the valid intervals. The requirement is then satisfied.

In a later step, this could also be integrated in a GUI framework, which automatically changes the plot of the raw R-R interval to a cleaned plot as suggested in sections 5.3.2, 5.3.3 and 5.3.4 after a cleaning of outlier intervals.

6.4 Usability Evaluation

The summary and display of the evaluation of the Polar H7 instrument by its participants is another requirement as explained in section 4.3 and presented in the User Interface Design in chapter 5. In order to achieve this in Python, a script was written importing first the csv module to allow the read-in of the participant data file containing the required columns with the answers given by the participants. After this, the specific columns were iterated through adding the values together and building an average value for each statement.
With regard to the requirement 8. of section 4.3, the result is expected to be:

The instructions on fitting the chest strap were clear:	3.666666
I found the chest strap easy to fit:	3.333333
Once fitted, the chest strap was comfortable to wear:	3.5
How long did it take you to fit the chest strap?:	1.5
How many attempt did it take you to fit the chest strap?:	1.333333

Figure 32: Usability Evaluation from participant data file

Result of testing the code: The executed code shows the expected output and the fulfilment of the requirement.

The range of answers for the first three statements were from 0 to 4 as explained in appendix 1.1 and have to be displayed later in the GUI as well to explain the participants' answers to the user. Nevertheless, the focus was set in this example on the operational functionality of the code to receive the summarised results, which can be easily integrated in a later GUI.

6.5 Polar Data File Pre-Processing

Another requirement for the software tool to be developed in Python is the pre-processing of the polar data file of each participant (requirement 3 in section 4.3). As mentioned in the beginning of this work, the Polar H7 instrument measures the R-R interval which is stored in a file and other HR, Distance/Speed measurements which are stored in another file (see polar data file appendix 1.3). Nevertheless, both are measured simultaneously.

However, as shown in appendix 1.3, there is no column which tags the data rows of the polar data file with the appropriate task instance that was conducted while measuring the different data. Nevertheless, this is important especially when analysing the heart rate variability of the task instances as discussed in section 5.2.3, in order to have additional data to integrate in the analysis process. Therefore, a script was written which matched the seconds of the R-R interval of a task instance with the time column in the polar data file, and tagged each row which was measured in the same time with its corresponding task instance.

To achieve this requirement, at first the csv module was imported to read in the polar data and the R-R interval file of a participant (task 3 instance 1 of nurse 1 was chosen as sample data as well as her polar data file). Furthermore, a daytime module was imported to manipulate the time of the polar data file (“00:00:00”) and convert it into seconds in order to be able to match it with the seconds of the R-R instance. Therefore, the time of the R-R interval and the time of the polar data file were compared to receive a matching time span. Accordingly, the rows in the polar data file which were recorded at the same time as the R-R interval were tagged with the task instance of the R-R interval file in new column.

```
## Set Task Definition
#Get first time value and last time value of R-R interval and iterate through polardata and identify matching time interval where task has to be set
def setTask(polarDataArray, heartbeatArray):
    firstTimeStamp = heartbeatArray[0].timeInSeconds -1  ### get first element
    lastTimeStamp = heartbeatArray[-1].timeInSeconds -1  #### get last element

    for i, polar in enumerate(polarDataArray):
        if (polar.timeInSeconds >= firstTimeStamp and polar.timeInSeconds < lastTimeStamp):
            polar.Task = "Task3/1"
```

Figure 33: Set task for matching time interval

Therefore, a new polar data file was written and put into an output folder in which the polar data rows are tagged to their appropriate task sequence.

An acceptance test was conducted to test the code with the above mentioned sample data of Nurse 3. The expected result was: R-R interval of Nurse 3, Task 1, Instance 1 starts at second **7941** and lasts until second **8060**. The matching interval of the polar data file starts in the time column at **02:12:21** = ((2h*60+12min) *60+21s=**7941s**) and lasts until **02:14:20** = ((2h*60+14min) *60+21s=**8060s**). each of these rows in the polar data file have to be tagged with Task3/1 to allocate the matching task instance to each row.

Result of testing the code: The executed code writes a new file in the output folder with a new column that tags these rows with Task3/1. The requirement is satisfied.

```
Task ID,Time,Time of day,HR [bpm],Speed [km/h],Distance [m],Acceleration [m/s□],Running cadence
```

Figure 34: Adding of Task ID to polar data file

6.6 Video Processing

The following sections discuss Python’s ability to deal with the video data of this project.

6.6.1 Anonymization of video material

As mentioned in section 2.3.2, privacy has to be protected in data analysis. Therefore, requirement 7. of the requirements mentions that before the recorded video can be analysed, the participants have to be anonymised. An often used practice is the blurring of the faces in the video. This can be done either manually (frame by frame), which may take a large amount of time or programmatically. However, with a look at the recorded videos of this study, the participants often turn around or can only be seen from the side. This might lead to risks that the face blurring is not 100% reliable and reveals the face of the person. Therefore, the idea arose to deface the complete image of the video. This means that the resolution of the video can be lowered to a certain degree so the person is not recognisable anymore, but all movements are still observable. This is important due to the discussed alignment of the graph and the video material for the analysis of the heart rate variability and outliers (see section 4.2.4 and section 5.2.3).

Python offers for the manipulation of image and video content the library openCV which has different possibilities for image and video processing (Bradski, 2000). Nevertheless, due to the short time of this project this requirement was not implemented with openCV. The goal was to find a method for anonymization of video content. Figure 33 shows two examples of how the resulting video could look like.

Figure 35: Negative effect on video (left) / blur effect on video (right)

6.6.2 Video Segment Split Python

As discussed above and with a look at the requirements 1. and 2. of section 4.3., the video of a task instance of a participant should be used in the analysis of the heart rate variability data. As mentioned in the design chapter of this work, the graph should be interactive, which means that if video material exists for a task instance, it has to be aligned with the graph (both should have the same recording time so it can be aligned over the seconds). Therefore, the user only clicks on a segment in the graph where for example, an ectopic interval exists and the video of this segment appears, showing the participant's action in that time. The alignment of the graph and the video material is not going to be discussed further in this section because this is going to be done within the development of the Graphical User interface of the tool in a future step.

However, python's ability to display the video for the chosen graph segment is shown. With regard to this, a script was written importing the os module and the moviepy editor to read the whole file and cut out the video segment needed to later save it in a new "sub" file.

```
###Input
timefocus = 872
duration = 2
inputFilename = "Chest Compressions.mp4"
outputFilename = "sub_" + inputFilename

###Processing
filepath = os.path.join(os.path.dirname(__file__), inputFilename)
video = VideoFileClip(filepath).subclip(timefocus - duration/2,timefocus + duration/2)

###Save file
video.write_videofile(outputFilename,fps=25)
```

Figure 36: Video Split in Python

With a look at an example of Nurse 1 and task 3 (see section 4.2.4, Figure 10), an ectopic interval can be detected in roughly second 86 of the graph. Therefore, if the user clicks on the ectopic interval to analyse its reason, normally the time focus is set to 86 and a duration can be predetermined on how long the video segment should be. The example above was tested with the video material available (chest compression video of cohort 1). Second 86 in the graph represents second 872 in the recorded video. A duration functionality was also built-in so that the displayed video segment can be modified in length. In this example, it was set to 2 seconds, so a video segment from second 871 to second 873 was expected, showing the movement of the participant during second 872.

Result of testing the code: The video saves a new file which is 2 seconds long, from second 871 to second 873. Test was successful.

As can be seen above, the existing recording of the video material is not ideal. Smaller files can be faster processed and can be easier aligned with the graph. Therefore, the video material should be pre-processed as well in unique files for each participant to avoid a second count to 872 as shown in the example above.

6.7 Calculation of Time-Domain Measures

The selected analysis method for the R-R interval for the software tool is the Time-domain method. Therefore, four statistical measures (Mean, SD, RMSSD, pNN50) were determined in the previous chapters, which can be applied to determine the heart rate variability of the R-R interval. A module for Heart Rate Variability analysis developed by Bartels (2015) could be found on github, providing these measurements so that they can be applied in Python for R-R interval data. The module uses the libraries numpy and scipy for the calculations, and furthermore provides a calculation possibility for the frequency-domain analysis as well. The evaluation of the software tool in section 5.3 revealed that an addition of frequency-domain analysis components might be beneficial for the further development of the system.
Due to the lack of time by the end of this project, the package wasn't applied and tested for the recorded data of the study. Nevertheless, the module can serve as a help for the further development of the software tool in Python.

6.8 Conclusion

The proof of concept of the examples provided above show that it is easy to write Python scripts with a simple text editor. This makes the code development more flexible and also easy to maintain. In regard to the analysis of wellbeing data, the different examples shown above validated Python's operational capability for processing the recorded R-R interval and visualising it. In addition to that, also a module exists for Python on github which provides scripts that are applying the time-domain analysis as well as frequency-domain analysis methods to recorded R-R interval data. In general, it can be said that structured data in tables and spreadsheets – which is dominant in this project, can be easily processed through Python. With regard to the unstructured data of this project, it was further discussed and proved that Python can be used to anonymise and fetch the required time segment from a video that wants to be reviewed during the analysis of the heart rate variability graphs. The acceptance tests confirmed the operational functionality of the code and the fit to the requirements to be later implemented in a GUI. All scripts as well as the results of the acceptance tests, can be reviewed in the digital appendix of this work (CD-ROM).

7 Evaluation

This chapter includes the evaluation of the dissertation project. The project management of the work, scope and achieved results are discussed as well as the research approach used. Furthermore, the results were compared to other studies in literature dealing with heart rate variability, as well as other software tools existing in the market for the analysis of heart rate variability.

7.1 Results

With regard to the project plan, all milestones could be achieved (see appendix 5). However, the implementation part could not be carried out as much as it was planned due to the lack of time by the end of the project. However, it has to be mentioned, that the first 2,5 weeks of the project were lost due to a shift in the topic. Therefore, the time of the project was shorter than planned. Also, the requirements and design chapter took longer than expected due to the detailed research on how a robust system for the analysis of wellbeing data should look like and which functionalities it should cover.

The scope of this work was the design of a system which combines the various types of data in this project and enables the user to answer arising questions about the data. In addition, the work had three research questions to answer:

- RQ1: What questions might be asked of the data?
- RQ2: Does the equipment (Polar H7) ensure a reliable data recording?
- RQ3: To what extent does the software tool enable research to draw conclusions about the data and provide a foundation for further investigations in this area?

With regard to the results, chapter 4, 5 and 6 could answer the research questions of this work. The Polar H7 instrument recorded two data files, the R-R interval and the HR and speed/distance measurements (polar data file, see appendix 1.3). First, the polar data file including the speed/distance measurements was evaluated over each participant, showing drop outs which were in the beginning and in the end of the recording period. A reason for this could be a poor skin contact especially in the starting period (when attaching it to the chest, dry skin etc.) and in the end of the recording time (before/when taking down the sensor). However, as discussed in section 2.1.2, when using body sensors there can be several reasons for errors in the data recording; so a precise reason couldn’t be given also due to the lack of video material during these periods where errors appeared. Furthermore, there also does not exist a standard in literature which indicates a “normal” outlier range, due to the mentioned variety of reasons for appearing inconsistencies in the data recording. Another error within the recorded data was the acceleration parameter which had a continuous 0 value over all the participants. Therefore, in this case, it can be assumed that this error might be caused by the device rather than the participants due to its existence in every recording. Without taking the acceleration parameter into account, the drop out values were minor with regard to the whole recording period.

With regard to the R-R interval which was recorded, as well by the Polar H7 instrument; data were recorded continuously over each participant. Nevertheless, outliers appearing in the recordings might have several reasons. Nevertheless, three analysed examples showed that movement artefacts seem to be one possible reason for these outliers. With regard to the data recording reliability, also talking the above findings of the additional recordings into account, it was concluded that the Polar H7 instrument showed a reliable data recording because it consistently records the R-R interval data and the HR, Speed/Distance measurements over all participants. However, with a look at the small population size of this study the data was received from, further research needs to be done as well with larger populations and possible comparisons with other recording instruments to further evaluate the reliability and also the validity of the recording of the Polar H7 sensor strap.

Apart from this, different studies in literature were reviewed which primarily used time-domain and frequency domain methods to analyse the HRV. The time-domain analysis method (in seconds and milliseconds in contrast to hertz in the frequency-domain) was chosen for this work in order to enable the user of the system so the graph of the R-R interval could be better aligned with other recorded data of the study (especially video and other polar recordings) during the analysis process (see section 5.2.3).

Three examples were conducted for Task 3 with participant 1, 2, and 3 to show a proof of concept on how the data can be analysed with the system in the future. Regarding the 181 task instances of this study, the data were not analysed over all tasks and participants. Apart from the fact that this was not the scope of the work, this also was not necessary due to the fact that the three examples were enough to show differences in the data which might be also expected from the rest of them, and raised questions to be answered as:

- How accurate is the Polar H7 instrument used in this study in regard to the measured R-R interval and the speed and distance measurements?
- What is the reason for the continuous 0 value for acceleration in the data, does a software error exist which is causing these phenomena?
- Is it possible to explain the reason for a high number of outliers in the recorded R-R interval of a participant?
    - With regard to participant 1 and 4, the outlier number was higher than in the recordings of the other 4 participants (see appendix 3.2).
        - Is it because the sensor did not fit properly?
        - Do participants who have a high number of outliers often touched the sensor during the recording time?
- How well does the video material confirm the reasons for outliers in the R-R interval besides the shown examples in this work?
- Is it possible to prevent the outliers in the R-R interval appearing through movement artefacts for the next study? (for instance, providing better instructions to the participants not to make abrupt upper body movements during the recording period)
- How well does the video material confirm the different experienced exposures and stress levels from the heart rate variability of the participants?
- How do the socio-demographic characteristics and health factors influence the heart rate variability? (comparison of different participant profiles).
- Are the data reproducible within an individual over longitudinal studies?

Followed by the evaluation of the data, requirements were created, and the system designed on their basis. The designed system was further evaluated by three experts in the field of wellbeing and healthcare through interviews and questionnaires. These data collection instruments could be regarded as valid, as they yielded similar results (Saunders et al., 2009). Furthermore the use of two collection instruments also enhanced the reliability of the received data as it can be seen in section 5.3 (Saunders et al., 2009; Sauro & Lewis, 2012). Overall the results of the evaluation were useful in order to improve the system in terms of design and features (see section 5.2 in comparison to section 5.3). However, a limitation of this work is that only three experts were interviewed and only one evaluation iteration was performed due to temporal limitations.

Nevertheless, with regard to research question three, the results from the interviews (see appendix 4.1) as well as the evaluation of the data in chapter 4 revealed to what extend the software could benefit further research in this area. First of all, with the integration of video material in the analysis process of the recorded R-R interval, reasons for inconsistencies could be found in order to improve further research studies in preventing outlier values caused by for instance, movement artefacts or consistent touching of the sensor by the participants. Furthermore, as mentioned by the interviewed experts, a better planning of shifts and tasks of nurses with regard to their exhaustion levels during the measured tasks could be performed in the future. This would offer the highest care potential to the patient. Furthermore, it was also mentioned that a real time monitoring of the heart rate variability and the individual's exhaustion and stress levels could be beneficial specifically in occupations with a high responsibility.

Furthermore, the system enables the intra subject comparability of the different task instances of a participant through a subject view (see section 5.3.3; Figure 20). This enables the user to analyse the varying heart rate variabilities of a participant over the same task over different recording periods. Furthermore, a task view (see section 5.3.4, Figure 21) is included in the system; which provides the possibility to compare the heart rate variability of all the studied individuals over the same task. In addition, the system highlights different socio-demographic and health factors which could benefit from drawing conclusions out of the data (f.e. if smoker was ticked, all smokers would be highlighted and the heart rate variability of smokers and non-smokers could be compared for a chosen task, see section 5.3.4).

Last but not least, the calculated time-domain measures can also be plotted and analysed, providing further information about the heart rate variability especially between the different task instances. An additional feature of the system is that the user can apply different cleaning parameter to the data, having the possibility to clean the data more moderately or aggressively.

The software architecture of the application revealed that a distributed code solution seems to be a beneficial foundation for the system. Especially with regard to Python, which was emphasised on this work, developers can easily access the different scripts written and maintain the code with a simple text editor because Python is not compiled.

As mentioned, in the final implementation of the required functionalities of the system, the focus in this work was set on Python, a multi-purpose programming language which enables fast prototyping and provides different tools for data analysis (see chapter 6). In addition to this, also packages for the build-up of a Graphical User Interface exist (see section 2.5). Nevertheless, the Graphical User Interface as designed in chapter 5 could not be implemented due to the lack of time by the end of the project. However, this was not needed to show a proof of concept of the key operational functionalities of the system in Python. These functionalities can be easily integrated in a GUI in a later step.

Therefore, the following table shows a summary of the sections where the designed and implemented results of this work can be found:

Table 6: Sections representing the implemented results of this work

Requirement number (summary of section 4.3 and 5.5)	Scripts (Proof of concept)	GUI
1	6.2 / 6.6.2	5.3.3
2	6.2 / 6.6.2	5.3.4
3	6.5	5.3.3
4	6.7	5.3.3 / 5.3.4
5	6.2	5.3.2
6	6.3	5.3.2 / 5.3.3 / 5.3.4
7	6.6.1	5.3.3 / 5.3.4
8	6.4	5.3.1

Acceptance tests proved the operational functionality of the implemented requirements and showed a proof of concept of the suitability of Python for the performance of the system design. As mentioned, due to the fact that Python scripts are not compiled, they can easily be written with a simple text editor and quickly maintained if for instance errors appear in the code, without the need of a development environment as eclipse in Java for instance. However, a GUI of the designed system was not implemented in this work. Further research needs to be done in this point also with regard to the later use of the system (through a web or desktop application), if Python should only be focused to perform the analytic side of the system or also be used for a build-up of the GUI.

With a critical evaluation of the research approach taken in this work, it can be concluded that the theoretical framework of the literature review around wellbeing data, data science and technology was a suitable starting point which provided a foundation for the further requirements and the design of the system. However, big data and data mining, as discussed in literature were not really the topic of this work. Nevertheless, this was still discussed as the scale up of the system through a web application could lead, for instance, to a variety of data over large sample sizes. Furthermore, also different mining algorithms as predictive analytic methods could be applied within a future development of the system. For example, the analysis over a large population could identify different health factors responsible for a low heart rate variability and a high amount of stress for a specific task. Therefore, a prediction could be given based on the socio-demographics and health factors of a nurse; for instance, what level of stress the individual might experience in a specific task.

Apart from that, the used qualitative research approach for the evaluation of the system design benefited the system and ensured it contains the core functionalities needed. On the other hand, there exist different limitations. One major limitation within qualitative research is the inexperience of the researcher as stated by Hussey and Hussey (1997). With regard to the authors, semi-structured interviews require certain skills to find the most appropriate answers for the research project. One of the reasons mentioned is the complexity in analysing the retrieved data and categorising emerging themes. Furthermore, in face-to face interviews the researcher always has a certain degree of influence on the interviewee which might lower the accuracy of the answers given by the respondent (Saunders et. al, 2009). Therefore, these reasons lower the reliability of the entire study.

Furthermore, the sample size of three interviewees is not representative as already mentioned in this work (Saunders et. al, 2009; Creswell, 2013; Sauro & Lewis, 2012). However, the quantitative questionnaire used as a back-up tool confirmed the repeatability of the answers given by the participants in the interviews and enhanced the reliability of the overall collected data.

The discussed built methodology used for the design and development of the system included a requirement, design, implementation and testing stage which was successfully followed. The only difference in this project is an evaluation stage after the design stage in order to assess the designed prototype system and adjust it to the demanded additional requirements received from the interviews. Acceptance Tests were used to check whether the implemented code satisfied the discussed requirements. Nevertheless, unit testing is also mentioned by different authors as a valid method and should be especially taken into consideration within the further development of the system; due to their strength in scanning chunks of code for identifying errors, which is crucial in the further development process of the system (Eckstein & Baumeister, 2004; R. Rogers, 2004).

7.2 Other Studies in Literature

As discussed above, for the system of this work, the time-domain analysis method was chosen due to the aim of aligning the different types of data especially when plotting the R-R intervals and analysing them in comparison with the video material (see section 4.2.4). By comparing this work with other studies in literature dealing with heart rate variability data, many studies use time-domain measures as well as frequency domain measures to receive more accurate indications about stress levels (Aubert et. al, 2003; Taelman et. al, 2009; McNarry & Lewis, 2012; Ito et. al, 2001). Therefore, the focus on the time-domain analysis method applied in the system without including the frequency domain measures, can be seen as a limitation.
Accordingly, a frequency domain method for analysing heart rate variability data should be added in further work.

Also with regard to the setup of the conducted study, a comparison with other studies in literature revealed that the recording time (in general 2 min recording intervals) is very short (see section 4.2.4). Most of the studies in literature had a recording time starting from 5 minutes to 24 hours or 36 hours to receive more significant results (Billman, 2011; Dekker et. al, 2000; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Furthermore, as mentioned in section 4.2.4, the task instances of the study had sometimes a varying time recording (some are 3-4 min long), or did not have the same order, which is a clear limitation for the comparability of the data (Aubert et. al, 2003; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996; Von Borell et. al, 2007). Another limitation of the data of this work compared with other authors is the lack of standardization of the exercise intensity between the nurses (McNarry & Lewis, 2012). This means participants were measured at similar intensity levels without taking into account varying fitness levels which could also be a reason for the resulting outliers in this work, through mechanical movement artefacts for instance (McNarry & Lewis, 2012; Tulppo, Mäkikallio, Seppänen, Laukkanen, & Huikuri, 1998).

Furthermore, results of the HRV analysis of three participants conducted in section 4.2.4 revealed differences among the participants which raised questions especially about the influence of the individual socio-demographics and health factors. In regard to other conducted studies, Van Amselvoort et. al, (2000) found out in their study that smokers had in general a lower heart rate variability than non-smokers, indicating the influence of living habits on the heart rate-variability. In addition to that, the study results of Järvelin-Pasanen et. al, (2012) also discussed that the demographic characteristics, such as the age, play an important role in decreasing HRV.

Van Amselvoort et. al, (2000) further concluded that shift workers had a lower heart rate variability than individuals with "normal" working hours. Especially in nursing occupations, shift work is very common (see chapter 1) and might have various influences as well on the received data of this study and further studies.

With a look at the concluded data recording reliability of the Polar H7 instrument, no other studies could be found in literature which are using explicitly the Polar H7 sensor strap for their recordings. However, in terms of recording reliability the results of this work match similar findings of the studies conducted by Vanderlei et. al, (2008) and Junqueira and Porto (2009); who compared another Polar instrument (Polar S810) with the conventional ECG. The instrument was seen as a reliable tool due to the fact among others, that it provided data recordings of R-R series at rest and exercise over various instances, as it is the case of this work.

7.3 Other Software Systems

With a look at other existing software tools which are dealing with heart rate variability data, one other Python-based tool could be identified. The gHRV is a free available software designed for clinical staff and researchers. It was developed in Python and allows the user to analyse the R-R interval through time-domain, frequency-domain methods and non-linear methods (Rodriguez-Martinez et al., 2010). The rMMSD and pNN50 time-domain measures used in this work are also included in the gHRV software. The outlier removal can be applied to the data manually or automatically in the time-domain method whereas an interpolation is used in the frequency domain analysis of the data. Furthermore, the measures used in each analysis method can be plotted, as implemented in this work (see section 5.3.3 and 5.3.4).

Another often stated tool in research is the Kubios HRV software, a free tool which includes time and frequency domain methods as well as non-linear methods (Tarvainen, Niskanen, Lipponen, Ranta-aho, & Karjalainen, 2013). One difference to this work's approach – apart from the lack of non-linear and frequency domain measures, is that the system has a bigger variety of time-domain measures, giving the user more information about the data. An outlier cleaning method is existing as well where the user can chose a threshold for intervals to be cleaned (Tarvainen et al., 2013). All in all, the user has different views that can be accessed, showing the results of different analysis methods applied to the R-R interval data. This also includes plots, showing the R-R interval over time as displayed in the software tool of this work, and other non-linear and frequency domain graphs.

There are also many other cross-platform and platform-specific software solutions for the analysis of heart rate variability which are not going to be analysed closer, because this would be beyond the scope of this work (Singh et. al, 2015; Tarvainen et. al, 2013). Nevertheless, in general, time domain and frequency domain methods are used in software tools, whereas some are also use non-linear analysis methods as pointcaré plots for instance (Singh, Bharti, & Engineering, 2015). However the plotting of the graphs of time and frequency domain features is commonly used (Altini, 2013; Tarvainen et. al, 2013). All systems have the aim to analyse the HRV and only vary in their measures and the visualisations used (Singh et al., 2015).

The biggest difference of all the reviewed tools with regard to the system of this work is, apart from the broader offer of analysis methods in these systems, the measures and visualisation opportunities of these tools on the market; that no system could be found that is using additional video material or other recordings (Speed/distance measurements, see appendix 1.3; Socio-demographics and health factors, see appendix 1.1) to align them with the analysis process of the heart rate variability and outliers in the data. Furthermore, all reviewed systems allow the analysis of only one data set of a participant at a time; whereas the system of this work allows for the analysis of multiple task instances in the subject view (see section 5.3.3) as well as the direct comparison with other participants' data through a task view (see section 5.3.4).

7.4 Conclusion

Overall, the project is a success. The work designed and evaluated a system for the analysis of wellbeing data. Furthermore, during the evaluation of the recorded data by the Polar H7 instrument and during the design and evaluation process of the system, all three research questions could be answered. The programming language Python was used in this work for the implementation of the system. Chapter 6 confirmed Python's flexibility to manipulate, analyse and visualise the recorded wellbeing data of this study. The proof of concept in chapter 6 showed, that Python can be used for the further development of the functionalities of the system.

In comparison with other systems on the market, it can be seen that commercial and free available tools have a more complex GUI with additional offered analysis methods and functionalities. Nevertheless, with regard to the aim of the system of this work to provide a foundation for research concerning the monitoring of the heart rate variability of the participants and detection of inconsistencies in the R-R interval data as outlier intervals; the implemented measures and visualisations of the system in this work provide an application easy to utilize by the user and clearly satisfy this aim. In addition to that, the prototype system of this work allows the comparison of multiple participants at a time, which has not been provided by any reviewed system on the market. Furthermore, none of the reviewed tools on the market integrated video data in the HRV analysis process, which is one of the strengths of this work's system. However, in later steps, other analysis methods and functionalities should be added on top of the system to compete with existing software on the market. On top of that, different other studies were reviewed, and suggestions for further research in this field were given on how the recording of the heart rate data could be further improved to receive more significant data which can be analysed with the system of this research project.

8 Conclusion and further research

This work designed a system to enable its users to efficiently analyse wellbeing data. The focus of the system is set on the analysis of heart rate variability data recorded from individuals during different task instance. Similar systems which analyse heart rate variability are available in the commercial market. Nevertheless, these tools differ from the system of this work in their analysis methods used and additional built-in features. However, this work introduced a system whose strength is to integrate different data to provide as much information as possible for the analysis of the heart rate variability data. This further lays a foundation for further research studies to draw conclusions and answer arising questions from the recorded data.

The requirements of the system were defined through an evaluation of the recorded data as well as by previous project meetings with the research team that recorded it. These requirements were used to design the software system of this work which was evaluated by three experts in the field of healthcare. The evaluation process revealed the usefulness of the designed system and further helped to improve and extend the functionality of it throughout this work.

In addition to this, during the evaluation process of the recorded data by the healthcare study, differences in the heart rate variability of the participants could be detected. Nevertheless, research is still ongoing, there are no results yet. However, a foundation was built with the system in this work, to further analyse the data, find patterns and answer the arisen questions.

The architecture of the system was designed and discussed. It showed the system’s strong foundation including a distributed folder structure where code is stored on various scripts so each script can be easily manipulated without leading to a total breakdown of the whole system. Especially with regard to the used programming language in this work, Python; the scripts are not compiled and can therefore easily be changed or extended. This means that a developer can write scripts with a simple text editor and store them in the system folder structure. This provides flexibility and an easy maintenance.

This work focused on Python due to its multiple strength in Data analysis and visualisation. As shown in chapter 6, key requirements of the system were implemented with Python, which can be perceived as appropriate for fulfilling the requirements of this work. Nevertheless, a GUI was not implemented. Further research should be done with regard to what tools are most suitable for the implementation of the GUI. The later environment where the system is used, should be taken into account as well. Apart from that, the algorithm discussed in section 4.2.4 as well as all functions for data validation, processing, and graph plotting should be implemented before implementing the more complex GUI. This approach would be good for a user with knowledge on how to run python scripts and would allow to analyse the data fast (see section 5.1.2). It would give a flavour on what is important for a further development of a GUI. Implementing the GUI directly would take more development time and decrease the flexibility for later changes.

References

Agusthiyar, R., & Narashiman, K. (2014). A Simplified Framework for Data Cleaning and Information Retrieval in Multiple Data Source Problems. *International Journal of Innovative Research in Science, Engineering and Technology*, *3*(8), 15401–15407.

Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Berger, A. C., Cohen, R. J., (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, *213*(4504), 220-222.

Albert, N. M., Butler, R., & Sorrell, J. (2014). Factors Related to Healthy Diet and Physical Activity in Hospital-Based Clinical Nurses. *Online journal of issues in nursing, 19*(3), 5.

Altini, M. (2013). Hear rate variability Logger – App details. Retrieved July 20, 2016, from: http://www.marcoaltini.com/blog/heart-rate-variability-logger-app-details.

Altini, M. (2016). *Issues in heart rate variability (HRV) analysis: motion artefacts & ectopic beats*. Retrieved 17 June, 2016 from: http://www.hrv4training.com/blog/issues-in-heart-rate-variability-hrv-analysis-motion-artifacts-ectopic-beats

Amaral, J., Buro, M., Elio, R., Hoover, J., Nikolaidis, I., Salavatipour, M., Steward, L., Wong, K. (2011). *About computing science research methodology*. Retrieved 20 July 2016 from: https://webdocs.cs.ualberta.ca/~c603/readings/research-methods.pdf.

Amtmann, B. J., & Amtmann, K. (2010). Use it or lose it : Physical fitness for nurses. *American Nurse Today*, *5*(July), 45–47.

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. S., & Buyya, R. (2015). Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79*, 3–15.

Atallah, M., Bertino, E., Elmagarmid, A., Ibrahim, M., & Verykios, V. (1999). Disclosure limitation of sensitive rules. *Proceedings of the 1999 Workshop on Knowledge and Data Engineering Exchange (KDEX'99) (Cat. No.PR00453),* 45–52. Retrieved 16 June, 2016 from: http://doi.org/10.1109/KDEX.1999.836532.

Aubert, A. E., Seps, B., & Beckers, F. (2003). Heart Rate Variability in Athletes. *Sports Med*, *33*(12), 889–919.

Baars, H., & Kemper, H.-G. (2008). Management Support with Structured and Unstructured Data—An Integrated Business Intelligence Framework. *Information Systems Management*, *25*(2), 132–148.

Bakshi, K. (2012). Considerations for big data: Architecture and approach. *IEEE Aerospace Conference Proceedings*, 1–7. Retrieved 02 July, 2016 from: doi: 10.1109/AERO.2012.6187357

Bartels, R. (2015). Module for Heart Rate Variability Analysis. Retrieved August 5, 2016, from: http://github.com/rhenanbartels/hrv.

Berntson, G., & Cacioppo, J. (2004). Heart rate variability: Stress and psychiatric conditions in: *Dynamic Electrocardiography* (Malik, M. & Camm, A. J., eds), Blackwell Publishing, Oxford, UK.

Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don’t miss a beat! *Psychophysiology*, *35*(1), 127–132.

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossmann, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H. & Van der Molen, M. W. (1997), Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6): 623–648.

Berry, M. J. A., & Linoff, G. (2004). *Data mining techniques : for marketing, sales, and customer relationship management*. Wiley.

Billman, G. E. (2011). Heart rate variability - A historical perspective. *Frontiers in Physiology*, *2*(86), 1–13.

Bishop, M., Cummins, J., Peisert, S., Singh, A., Bhumiratana, B., Agarwal, D., … Hogarth, M. (2010). Relationships and data sanitization. In *Proceedings of the 2010 workshop on New security paradigms - NSPW ’10*, 151-164. Retrieved 19 June, 2016 from: http://doi.org/10.1145/1900546.1900567.

Blake, H., Mo, P. K., Lee, S., & Batt, M. E. (2012). Health in the NHS: lifestyle behaviours of hospital employees. *Perspectives in Public Health*, *132*(5), 213–215.

Boullosa, D. A., Abreu, L., Tuimil, J. L., & Leicht, A. S. (2012). Impact of a soccer match on the cardiac autonomic control of referees. *European Journal of Applied Physiology*, *112*(6), 2233–2242.

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication and Society*, *15*(5), 662–679.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools, 25*(11), 120-126.

Buchheit, M. (2014). Monitoring training status with HR measures: do all roads lead to Rome?. *Frontiers in physiology*, *5*(73), 1-19.

Buchheit, M., Mendez-Villanueva, A., Quod, M. J., Poulos, N., & Bourdon, P. (2010). Determinants of the variability of heart rate measures during a competitive period in young soccer players. *European Journal of Applied Physiology*, *109*(5), 869–878.

Burton-Jones, A., & Hubona, G. S. (2005). Individual differences and usage behavior. *ACM SIGMIS Database*, *36*(2), 58–77.

Chandola, T., Britton, A., Brunner, E., Hemingway, H., Malik, M., Kumari, M., Badrick, E., Kivimaki, M. & Marmot, M. (2008). Work stress and coronary heart disease: what are the mechanisms? *European Heart Journal*, *29*(5), 640–648.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209.

Cios, K., Pedrycz, W., Swiniarski, R., & Kurgan, L. (2010). *Data mining: a knowledge discovery approach*. Retrieved 15 June 2016 from: http://dl.acm.org/citation.cfm?id=1941721.

Clifford, G. (2006). Ectopic Beats , Activity Effects and Heart Rate Turbulence. Retrieved 8 June 2016 from: https://www.physionet.org/events/hrv-2006/clifford.pdf.

Creswell, J. W. (2013). *Research Design-Qualitative, Quantitative, and Mixed Methods Approaches* (4th edition). London: SAGE Publications Ltd.

Daliot, A. (2013). *Instant Axure Rp Starter*. Birmingham: Packt Publishing.

Dasseni, E., Verykios, V. S., Elmagarmid, A. K., & Bertino, E. (2001). Hiding Association Rules by Using Confidence and Support, 369–383, Springer Berlin Heidelberg.

Davis F. D. (1993). User acceptance of information technology: System characteristics user perceptions and behavioral impacts. *Internat. J. Man-Machine Stud.*, *38*(3), 475–487.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340.

Dekker, J. M., Crow, R. S., Folsom, a. R., Hannan, P. J., Liao, D., Swenne, C. a., & Schouten, E. G. (2000). Low Heart Rate Variability in a 2-Minute Rhythm Strip Predicts Risk of Coronary Heart Disease and Mortality From Several Causes : The ARIC Study. *Circulation*, *102*(11), 1239–1244.

Dhar, V. (2012). Data Science and Prediction. *Communications of the ACM*, *56*(12), 64–73.

Eckstein, J., & Baumeister, H. (2004). *Extreme Programming and Agile Processes in Software Engineering*. Garmisch-Partenkirchen: Springer.

EKG Interpretation. (2016). Arrhythmias. Retrieved 12 July, 2016 from: https://www.nurseslearning.com/courses/nrp/nrp1619/section%202/p01.html.

Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys*, *45*(1), 1–34.

Familoni, O. B. (2008). An overview of stress in medical practice. *African Health Sciences*, *8*(1), 6–7.

Fan, W., & Bifet, A. (2013). Mining Big Data : Current Status , and Forecast to the Future. *ACM SIGKDD Explorations Newsletter*, *14*(2), 1–5.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, *17*(3), 37-54.

Fekete, J.-D., Wijk, J. J., Stasko, J. T., & North, C. (2008). *The Value of Information Visualization BT - Information Visualization*. *Information Visualization Human-Centered Issues and Perspectives*. Berlin Heidelberg New York: Springer.

Fu, T. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, *24*(1), 164–181.

Fung, B. C. M., Wang, K., Chen, R., & Yu, P. S. (2010). Privacy-preserving data publishing. *ACM Computing Surveys*, *42*(4), 1–53.

Gamelin, F.-X., Baquet, G., Berthoin, S., & Bosquet, L. (2008). Validity of the Polar

S810 to Measure R-R Intervals in Children. *International Journal of Sports Medicine*, *29*(2), 134–138.

Gelsema, T., Van der Doef, M., Maes, S., Janssen, M., Akerboom, S. & Verhoeven, C. (2006). A longitudinal study of job stress in the nursing profession: causes and consequences. *Journal of Nursing Management, 14*, 289–299.

Giles, D., Draper, N., & Neil, W. (2016). Validity of the Polar V800 heart rate monitor to measure RR intervals at rest. *European Journal of Applied Physiology*, *116*(3), 563–571.

Goli-Malekabadi, Z., Sargolzaei-Javan, M., & Akbari, M. K. (2016). An effective model for store and retrieve big health data in cloud computing. *Computer Methods and Programs in Biomedicine*, *132*, 75–82.

Graham, D., Van Veenendaal, E., & Evans, I. (2008). *Foundations of Software Testing: ISTQB Certification*. London: Cengage Learning EMEA.

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., Duke C. S. & Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment*, *11*(3), 156–162.

Hand, D., Hand, D., Mannila, H., Mannila, H., Smyth, P., & Smyth, P. (2001). Principles of data mining, *Drug safety,30*(7), 621-622.

Haynes, G. R., & Lisicki, A. (1996). Unique cause for 60-cycle interference. *Anesthesiology*, *85*(5), 1215–1216.

Hey, T., Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: data-intensive scientific discovery*. Redmond, WA: Microsoft research.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, *9*(3), 90–95.

Huntington, A., Gilmour, J., Schluter, P., Tuckett, A., Bogossian, F. & Turner, C. (2009). The Internet as a research site: establishment of a web-based longitudinal study of the nursing and midwifery workforce in three countries. *Journal of Advanced Nursing, 65*(6), 1309–1317.

Hoffer, J., George, J. F., & Valacich, J. (2005). Modern Systems Analysis and Design (Fourth Edition). Prentice Hall.

Hongoro, C., & McPake, B. (2004). How to bridge the gap in human resources for health. *The Lancet*, *364*(9443), 1451–1456.

Hsu, M. K., Wang, S. W., & Chiu, K. K. (2009). Computer attitude, statistics anxiety and self-efficacy on statistical software adoption behavior: An empirical study of online MBA learners. *Computers in Human Behavior*, *25*(2), 412–420.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, *9*(3), 90–95.

Intel IT Center, & Intel. (2012). Big Data Analytics. Intel's IT Manager Survey on How Organizations Are Using Big Data. *Intel IT Center*, (August 2012). Retrieved 21 June, 2016 from: http://www.intel.com/content/dam/www/public/us/en/documents/reports/big-data-analytics-2013-peer-research-report.pdf.

Ito, H., Nozaki, M., Maruyama, T., Kaji, Y., & Tsuda, Y. (2001). Shift work modifies the circadian patterns of heart rate variability in nurses. *International journal of cardiology, 79*(2), 231-236.

Jagtap, S. B., Kodge, B. G. (2013). Census Data Mining and Data Analysis using WEKA. *(ICETSTM – 2013) International Conference in "Emerging Trends in Science, Technology and Management-2013, Singapore*, 35–40. Retrieved 20 June, 2016 from: https://www.researchgate.net/profile/Bheemashankar_Kodge/publication/257882673_Census_Data_Mining_and_Data_Analysis_using_WEKA/links/56d51f8108aefd177b0f667f.pdf.

Java. (2016). How can I get started developing Java programs with the Java Development Kit (JDK)? Retrieved July 20, 2016, from http://java.com/en/download/faq/develop.xml.

Junqueira, L. F., & Porto, L. G. G. (2009). Comparison of time-domain short-term heart interval variability analysis using a wrist-worn heart rate monitor and the conventional electrocardiogram. *Pacing Clin Electrophysiol*, *32*(2009), 43–51.

Kahn, L., & Cannell, C. (1957). *The dynamics of interviewing: theory, technique, and cases*. New York: John Wiley & Sons.

Kaisler, S., Armour, F., Espinosa, J. a, & Money, W. (2013). Big Data: Issues and Challenges Moving Forward. *46th Hawaii International Conference on System Sciences (HICSS)*, 995–1004.

Kakunje, A., & Professor, A. (2011). Editorial: Stress Among Health Care Professionals -The Need for Resiliency. *Online J Health Allied Scs*, *10*(1), 1.

Kamath, M. V., Watanabe, M. A., & Upton, A. R. M. (2013). *Heart Rate Variability (HRV) Signal Analysis: Clinical Applications*. Boca Raton: CRC Press.

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), 2561–2573.

Karlsson, M., Hörnsten, R., Rydberg, A., & Wiklund, U. (2012). Automatic filtering of outliers in RR intervals before analysis of heart rate variability in Holter recordings: a comparison with carefully edited data. *BioMedical Engineering OnLine*, *11*(1), 1-12.

Keim, D., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. In *Visual Data Mining*, (Simoff, S.J., Böhlen, M. H. & Mazeika, A., eds.), 76–90, Heidelberg, Germany: Springer.

Kemper, K. J., Hamilton, C., & Atkinson, M. (2007). Heart Rate Variability: Impact of Differences in Outlier Identification and Management Strategies on Common Measures in Three Clinical Populations. *Pediatric Research*, *62*(3), 337–342.

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., Shiraz, M. & Gani, A. (2014). Big data: Survey, technologies, opportunities, and challenges. *Scientific World Journal*, *2014*, 1–18.

Kim, D., & Chang, H. (2007). Key functional characteristics in designing and operating health information websites for user satisfaction: An application of the extended technology acceptance model. *International Journal of Medical Informatics*, *76*(11), 790–800.

Kleiger, R. E., Miller, J. P., Krone, R. J., & Bigger, J. T. (1990). The Independence of Cycle Length Variability and Exercise Testing on Predicting Mortality of Patients Surviving Acute Myocardial Infarction. *The American Journal of Cardiology*, *65*(7), 408–411.

Kohavi, R. (2000). Data Mining and Visualization. *Science*, *40*(4), 438–47.

Konold, T., Bone, G. E., & Simmons, M. M. (2011). Time and frequency domain analysis of heart rate variability in cattle affected by bovine spongiform encephalopathy. *BMC Research Notes, 4*(1), 1-12.

Koukis, D., Antonatos, S., Antoniades, D., Markatos, E. P., & Trimintzios, P. (2006). A Generic Anonymization Framework for Network Traffic. In *2006 IEEE International Conference on Communications, 5,* 2302–2309.

Lalani, V., & Meniya, A. (2015). Bigdata Analytics Using RESTful Web Services, *Journal of Emerging Technologies and Innovative Research, 2*(4), 954–957.

Lippman, N., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology, 267*(1), H411-H418.

Lopes, P. & White, J. (2006). Heart Rate Variability Measurements and Practical Implications, in: *Physiological Assessment of Human Fitness* (Maud Peter J., Foster, C., eds), 39-62, Human Kinetics.

Luckow, A., Santcroos, M., Zebrowski, A., & Jha, S. (2015). Pilot-Data: An abstraction for distributed data. *Journal of Parallel and Distributed Computing, 79-80*, 16–30.

Maciel, B. C., Gallo, L., Marin Neto, J. A., Lima Filho, E. C., & Martins, L. E. (1986). Autonomic nervous control of the heart rate during dynamic exercise in normal man. *Clinical Science*, *71*(4), 457–460.

Mann, B., Williams, R., Atkinson, M., Brodlie, K., Storkey, A., & Williams, C. (2002, October). *Scientific data mining, integration, and visualization*. In Report of workshop held at the National e-Science Institute. Retrieved 20 June, 2016 from: https://core.ac.uk/download/files/200/4892835.pdf.

Martin, R. C. (2003). *Agile software development: principles, patterns, and practices*. Prentice Hall PTR.

Mateo, J., & Laguna, P. (2003). Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Biomedical Engineering, 50*(3), 334-343.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data. The management revolution. *Harvard Business Review, 90*(10), 61-67.

McCallister, E., Grance, T., & Scarfone, K. (2010). *Guide to protecting the confidentiality of personally identifiable information. Retrieved 20 June, 2016 from:* http://doi.org/10.6028/NIST.SP.800-122.

McClure, S. (2015). Data Science and Big Data: Two very Different Beasts. Retrieved June 5, 2016, from http://www.kdnuggets.com/2015/07/data-science-big-data-different-beasts.html.

McKinney, W. (2013). *Python for data analysis*. Sebastopol, CA: O'Reilly Media.

McNarry, M. A., & Lewis, M. J. (2012). Heart rate variability reproducibility during exercise. *Physiological Measurement*, *33*(7), 1123–1133.

McNeely, E. (2005). The consequences of job stress for nurses' health: Time for a check-up. *Nursing Outlook*, *53*(6), 291–299.

Michael, K., & Miller, K. (2013). Big data: New opportunities and new challenges. *Computer*, *46*(6), 22–24.

Navarro, E., & Schrauf, R. W. (2005). Using Existing Tests and Scales in the Field. *Field Methods*, *17*(4), 373–393.

Nelson, G. S. (2015). Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification. Retrieved 15 June, 2016 from: thotwave.com/wp-content/uploads/2015/09/data_sharing_privacy_anonymization_and_de-identification_rev_13.pdf

Oliveira, S. R. M., & Zaane, O. R. (2003). Protecting Sensitive Knowledge By Data Sanitization, *Proc. 3rd IEEE International Conference Data Mining* (ICDM 03), 613–616.

Padilla-Meléndez, A., del Aguila-Obra, A. R., & Garrido-Moreno, A. (2013). Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario. *Computers & Education*, *63*, 306–317.

Park, J., Lee, S., & Jeon, M. (2009). Atrial fibrillation detection by heart rate variability in Poincare plot. *Biomedical engineering online, 8*(1), 1.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Thousand Oaks, CA, Sage.

Perez, F., & Granger, B. E. (2007). IPython: A System for Interactive Scientific Computing. *Computing in Science & Engineering*, *9*(3), 21–29.

Pipilis, A., Flather, M., Ormerod, O., & Sleight, P. (1991). Heart rate variability in acute myocardial infarction and its association with infarct site and clinical course. *The American Journal of Cardiology*, *67*(13), 1137–1139.

Polar. (2016). H7 heart rate sensor. Retrieved June 17, 2016, from:

http://www.polar.com/en/products/accessories/H7_heart_rate_sensor.

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Data Science and Big Data*, *1*(1), 51–59.

Python (2016). GUI Programming in Python. Retrieved 8 June, 2016 from: http://doi.org/https://wiki.python.org/moin/GuiProgramming.

Rahm, E., & Do, H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, *23*(4), 3–13.

Rodriguez-Martinez, M., Seguel, J., & Greer, M. (2010). Open source Cloud Computing tools: A case study with a weather application. *Proceedings - 2010 IEEE 3rd International Conference on Cloud Computing, CLOUD 2010*, 443–449. Retrieved 22 June, 2016 from: doi: 10.1109/CLOUD.2010.81.

Rogers, E. (2010). *Diffusion of Innovations* (4th ed.). New York. Simon and Schuster.

Rogers, R. O. (2004, August). *Acceptance testing vs. unit testing: A developer's perspective.* In Conference on Extreme Programming and Agile Methods (Zannier, C., Erdogmus, H. & Lindstrom, L., eds), 22-31, Berlin, Germany: Springer.

Russom, B. P. (2011). *Big Data Analytics. TDWI Best Practices Report, Fourth Quarter*. Retrieved 22 June, 2016 from: http://www.tableau.com/sites/default/files/whitepapers/tdwi_bpreport_q411_big_data_analytics_tableau.pdf

Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *International Conference on Collaboration Technologies and Systems (CTS), 2013 International Conference*, 42–47, San Diego, CA. Retrieved from: doi: 10.1109/CTS.2013.6567202.

Saunders, M., & Lewis, P. (2012). *Doing research in business and management: An essential guide to planning your project*. Harlow, Essex: Pearson Eductation.

Saunders, M., & Lewis, P. & Thornhill, A. (2009). *Research Methods for Business Students (Fifth edition)*. Milan, Italy: Pearson Eductation.

Sauro, J., & Lewis, J. R. (2012). *Quantifying the User Experience: Practical Statistics for User Research*. Elsevier.

Schluter, P. J., Turner, C., Huntington, A. D., Bain, C. J., & McClure, R. J. (2011). Work/life balance and health: the Nurses and Midwives e-cohort study.

International Nursing Review, *58*(1), 28–36.

Sharma, E. (2005). Role Stress Among Doctors. *Journal of Health Management*, *7*(1), 151–156.

Shmatikov, P., & Narayanan, A. (2010). Myths and Fallacies of "Personally Identifiable Information." *Communications of the Academic*, *53*(6), 22–26.

Singh, B., & Bharti, N. (2015). Software tools for heart rate variability analysis. *International Journal of Recent Scientific Research*, *6*(4), 3501-3506.

Solem, K., Laguna, P., & Sornmo, L. (2006). An efficient method for handling ectopic beats using the heart timing signal. *IEEE transactions on biomedical engineering, 53*(1), 13-20.

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542.

Swan, M. (2015). Philosophy of big data: Expanding the human-data relation with big data science services. *Proceedings - 2015 IEEE 1st International Conference on Big Data Computing Service and Applications, BigDataService 2015*, 468–477. Retrieved 19 June, 2016 from: http://doi.org/10.1109/BigDataService.2015.29.

Sweetwater. (2011). *Stress and Heart Rate Variability*, 1–4, Retrieved 05 June, 2016 from: http://www.sweetwaterhrv.com/documentation/stressandhrv.pdf.

Sweetwater. (2013). *Heart Rate Variability 101. An Introduction to HRB and Interpreting Sweet Beat HRV Graphs and Numbers*, 1–16. Retrieved 05 June, 2016 from: http://www.sweetwaterhrv.com/documentation/HRV_Measurement_Explanation.pdf.

Taelman, J., Vandeput, S., Spaepen, a, & Huffel, S. Van. (2009). Influence of Mental Stress on Heart Rate and Heart Rate Variability. *Heart*, *29*(1), 1366–1369.

Tanita. (2016). *Understanding your Measurements If you can measure it, you can manage it!* Retrieved June 16, 2016, from: http://tanita.eu/tanita-academy/understanding-your-measurements.

Tarvainen, M. P., Niskanen, J., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2013). Kubios HRV – Heart rate variability. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–220.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17*, 354–381.

Tonello, L., Rodrigues, F. B., Souza, J. W. S., Campbell, C. S. G., Leicht, A., & Boullosa, D. A. (2014). The role of physical activity and heart rate variability for the control of work related stress. *Frontiers in Physiology*, *5*(67),1–9.

Trnka, A. (2014). Big data analysis. *European Journal of Science and Theology*, *10*, 143–148.

Tsuji, H., Larson, M. G. M., Venditti, F. J. F. J., Manders, E. E. S., Evans, J. J. C., Feldman, C. L. C., & Levy, D. (1996). Impact of Reduced Heart Rate Variability on Risk for Cardiac Events, *Circulation, 94*(11), 2850–2855.

Tulppo, M. P., Mäkikallio, T. H., Seppänen, T., Laukkanen, R. T., & Huikuri, H. V. (1998). Vagal modulation of heart rate during exercise: effects of age and physical fitness. *American Journal of Physiology-Heart and Circulatory Physiology, 274*(2), H424-H429.

Umarji, M., & Seaman, C. (2005). Predicting acceptance of software process improvement. In *Proceedings of the 2005 Workshop on Human and Social Factors of Software Engineering,* ACM, 1–6, St. Louis, Missouri. Retrieved 28 June, 2016 from: http://doi.org/10.1016/j.im.2013.12.003.

Uusitalo, A., Mets, T., Martinmäki, K., Mauno, S., Kinnunen, U., & Rusko, H. (2011). Heart rate variability related to effort at work. *Applied Ergonomics*, *42*(6), 830–838.

Van Amelsvoort, L. G. P. M., Schouten, E. G., Maan, A. C., Swenne, C. A., & Kok, F. J. (2000). Occupational determinants of heart rate variability. *International Archives of Occupational and Environmental Health, 73*(4), 255–262.

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science and Engineering*, *13*(2), 22–30.

Vanderlei, L. C. M., Silva, R. A., Pastre, C. M., Azevedo, F. M., & Godoy, M. F. (2008). Comparison of the Polar S810i monitor and the ECG for the analysis of heart rate variability in the time and frequency domains. *Brazilian Journal of Medical and*

Biological Research, *41*(10), 854–859.

Varian, H. R. (2014). Big Data: New Tricks for Econometrics †. *Journal of Economic Perspectives*, *28*(2), 3–28.

Venkatesh, V., Walton, S. M., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance use of Information Technology: Extending the unififed theory of acceptance and Use of Technology. *Forthcoming in MIS Quarterly*, *36*(1), 157–178.

Videman, T., Ojajarvi, A., Riihimaki, H., & Troup, J. D. G. (2005). Low Back Pain Among Nurses. *Spine*, *30*(20), 2334–2341.

Villars, R. L., & Olofson, C. W. (2014). White Paper Big Data : *What It Is and Why You Should Care Information Everywhere, But Where' s the Knowledge?* White Paper, IDC. Retrieved 16 June, 2016 from: http://www.tracemyflows.com/uploads/big_data/idc_amd_big_data_whitepaper.pdf

Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero, M., Mohr, E. & Prunier, A. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—a review. *Physiology & Behavior, 92*(3), 293-316.

Wallace, L. G., & Sheetz, S. D. (2014). The adoption of software measures: A technology acceptance model (TAM) perspective. *Information & Management*, *51*, 249–259.

Wu, H., Salzberg, B., Sharp, G. C., Jiang, S. B., Shirato, H., & Kaeli, D. (2005). Subsequence matching on structured time series data. *In Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, 682-693. Retrieved 29 June, 2016 from: doi: 10.1145/1066157.1066235.

Wu, X., Zhu, X., Wu, G.-Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, *26*(1), 97–107.

Yarbrough, A. K., & Smith, T. B. (2007). Technology acceptance among physicians: A new take on TAM. *Medical care research and review, 64*(6), 650-672.

Yarnell, J. (2008). Stress at work--an independent risk factor for coronary heart disease? *European Heart Journal, 29*(5), 579–580.

Yuan, S.C, Chou, M.C., Hwu, L.J, Chang, Y.O., Hsu, W.H., Kuo, H.W. (2009). An intervention program to promote health related physical fitness in nurses. *Journal of Clinical Nursing, 18*, 1404-1411.

Zapka, J. M., Lemon, S. C., Magner, R. P., & Hale, J. (2009). Lifestyle behaviours and weight among hospital-based nurses. *Journal of Nursing Management*, *17*(7), 853–860.

Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data.* McGraw-Hill Osborne Media.

Appendix

1. Overview of the recorded data (see digital appendix, CD-ROM)

1.1 Participant data file

Parameter	Explanation
Participant	1, 2, 3, 4, 5, 6
Age	• 1 (18-24) • 2 (25-29) • 3 (30-39) • 4 (40-49) • 5 (50-59)
Weight	In kg
BMI	BMI index (kg/m²)
Gender	• 1 (male) • 2 (female)
Field of practice	• 1 (Adult) • 2 (Child) • 3 (Learning Disability) • 4 (Mental health) • 5 (Midwifery)
Study year	Study Year number
Do you take part in physical activity to exercise on most days of the week for 30 minutes or more each time?	• 1 (Yes) • 2 (No)
Are you currently smoking?	• 1 (Yes) • 2 (No)
Do you eat five servings of fruit/vegetables a day?	• 1 (Rarely) • 2 (Sometimes) • 3 (Everyday)
Do you eat food high in fat and sugar?	• 1 (A few times a week) • 2 (Once a day) • 3 (2-3 times a day)
How often do you drink 8 glasses (2 litres) of water a day?	• 1 (Rarely) • 2 (Sometimes) • 3 (Everyday)
How often do you have a standard drink containing alcohol?	• 1 (Never) • 2 (Monthly or less) • 3 (2-4 times a month) • 4 (2-3 times a week) • 5 (4 or more times a week)
How many standard drinks containing alcohol do you have on a typical day when drinking?	• 1 (1-2) • 2 (3-4) • 3 (5-6) • 4 (7-9) • 5 (10 or more)
BodyFat	In percent
BodyWater	In percent
MuscleMass	In percent

DCI_BMR	The minimum level of energy the body needs per day at rest in order to function effectively (in calories)
MetabolicAge	In years
BoneMass	In kg
VisceralFat	In percent
The instructions on fitting the chest strap were clear.	• 0 (Strongly Disagree) • 1 (Disagree) • 2 (Neither agree or disagree) • 3 (Agree) • 4 (Strongly Agree)
I found the chest strap easy to fit.	• 0 (Strongly Disagree) • 1 (Disagree) • 2 (Neither agree or disagree) • 3 (Agree) • 4 (Strongly Agree)
Once fitted, the chest strap was comfortable to wear.	• 0 (Strongly Disagree) • 1 (Disagree) • 2 (Neither agree or disagree) • 3 (Agree) • 4 (Strongly Agree)
How long did it take you to fit the chest strap?	• 1 (less than 1 minute) • 2 (2-3 minutes) • 3 (4-5 minutes) • 4 (6mins or more)
How many attempt did it take you to fit the chest strap?	Number of attempts

The file consists the socio-demographics and health factors of each participant. It was converted from SPSS to a csv file, containing numerical descriptions for the parameters. Therefore, this table should give an overview about the meaning of the numbers for each parameter. The marked parameters in bold were measured with Tanita scales.

1.2 R-R interval

Nurse_01_R-R.txt

```
0.0 0.309
0.309 0.409
0.718 0.324
1.042 0.282
1.324 0.251
1.575 0.487
2.062 0.732
2.794 0.407
3.201 0.477
3.678 0.602
4.28 0.286
4.566 0.363
4.929 0.414
5.343 0.711
6.054 0.255
6.309 0.509
6.818 0.351
7.169 0.405
7.574 0.252
7.826 1.122
8.948 0.599
9.547 0.401
9.948 0.416
10.364 0.407
10.771 1.019
11.79 0.76
```

Figure 37: R-R interval

R-R interval: Left column showing the relative time the sample was taken (starting from 0.0) and the right column is the time in milliseconds for the R-R interval.

1.3 Recorded data by the Polar H7 beside the R-R interval

Time	Time of day	HR [bpm]	Speed [km/h]	Distance [m]	Acceleration [m/s²]	Running cadence
0:00:00	10:40:39					
0:00:01	10:40:40	131	0.0		0.0	0
0:00:02	10:40:41	135	1.0	0.0	0.0	0
0:00:03	10:40:42	135	4.6	0.0	0.0	30
0:00:04	10:40:43	136	4.4	1.3	0.0	59
0:00:05	10:40:44	136	3.9	2.4	0.0	57
0:00:06	10:40:45	137	3.7	3.5	0.0	54
0:00:07	10:40:46	137	4.1	4.5	0.0	58
0:00:08	10:40:47	141	4.1	5.7	0.0	57
0:00:09	10:40:48	141	3.9	6.8	0.0	57
0:00:10	10:40:49	141	0.8	7.8	0.0	57
0:00:11	10:40:50	142	0.6	7.8	0.0	0
0:00:12	10:40:51	143	0.0	8.2	0.0	0
0:00:13	10:40:52	143	0.0	8.2	0.0	0
0:00:14	10:40:53	138	0.0	8.2	0.0	0
0:00:15	10:40:54	138	0.0	8.2	0.0	0
0:00:16	10:40:55	138	0.0	8.2	0.0	0
0:00:17	10:40:56	138	0.0	8.2	0.0	0
0:00:18	10:40:57	138	0.0	8.2	0.0	0

Figure 38: Polar data file

- Time column: Shows how long the instrument is already recording.
- Time of day column: Shows the time of the day.
- HR[bpm] column: Shows the estimated heart rate per minute.
- Speed [km/h] column: Shows the current speed of the participant.

- Distance [m] column: Shows the current speed of the participant.
- Acceleration [m/s²] column: Shows the current acceleration of the participant.
- Running cadence column: Calculates the number steps taken by the participant per minute.

1.4 Time-table of executed tasks of all participants

A	B	C	D	E	F	G	H	I	J	K	L
Nurse_1_Time	Exercise	Nurse_2_Time	Exercise	Nurse_3_Time	Exercise	Nurse_4_Time	Exercise	Nurse_5_Time	Exercise	Nurse_6_Time	Exercise
12.23	2.00	12.23	1.00	12.22	6.00	10.55	1.00	10.55	2.00	10.55	3.00
12.25	6.00	12.25	2.00	12.25	1.00	10.58	2.00	10.57	3.00	10.58	4.00
12.28	1.00	12.28	6.00	12.27	2.00	11.01	3.00	11.00	4.00	10.59	5.00
12.31	2.00	12.31	7.00	12.30	6.00	11.03	4.00	11.00	5.00	11.00	6.00
12.33	6.00	12.31	8.00	12.33	7.00	11.03	5.00	11.03	6.00	11.03	7.00
12.35	7.00	12.33	1.00	12.33	8.00	11.05	6.00	11.05	7.00	11.03	8.00
12.35	8.00	12.35	2.00	12.36	1.00	11.08	7.00	11.05	8.00	11.05	9.00
12.38	1.00	12.38	6.00	12.39	2.00	11.08	8.00	11.07	9.00	11.07	1.00
12.41	2.00	12.41	7.00	12.40	6.00	11.10	9.00	11.10	1.00	11.10	2.00
12.43	9.00	12.41	8.00	12.44	9.00	11.13	1.00	11.11	2.00	11.13	3.00
12.47	1.00	12.43	9.00	12.47	3.00	11.15	2.00	11.15	3.00	11.15	4.00
12.50	2.00	12.47	2.00	12.50	4.00	11.18	3.00	11.17	4.00	11.15	5.00
12.53	3.00	12.50	3.00	12.50	5.00	11.20	4.00	11.17	5.00	11.17	6.00
12.55	4.00	12.52	4.00	12.52	6.00	11.20	5.00	11.19	6.00	11.19	7.00
12.55	5.00	12.52	5.00	12.55	7.00	11.23	6.00	11.23	7.00	11.19	8.00
12.57	6.00	12.54	6.00	12.55	8.00	11.29	7.00	11.23	8.00	11.22	9.00
13.00	7.00	12.59	7.00	12.58	9.00	11.29	8.00	11.24	9.00	11.24	1.00
13.00	8.00	12.59	8.00	13.00	1.00	11.27	9.00	11.27	1.00	11.27	2.00
13.02	9.00	13.02	9.00	13.02	2.00	11.30	1.00	11.30	2.00	11.29	3.00
13.05	1.00	13.04	1.00	13.05	3.00	11.32	2.00	11.33	3.00	11.32	4.00
13.07	2.00	13.06	2.00	13.08	4.00	11.35	3.00	11.34	4.00	11.32	5.00
13.10	3.00	13.08	3.00	13.08	5.00	11.37	4.00	11.34	5.00	11.34	6.00

Figure 39: Activities file

The activity timetable includes for each participant two columns; one column which shows the time of the day a task was conducted by a participant and an exercise column which contains the task number which was done at that time.

There are 9 exercises which were conducted by the participants:

(1) - Manoeuver a bed table or locker
(2) - Turn a manikin from one side to the other on a bed
(3) - Simulated Chest compression
(4) - Assist a researcher to transfer from a bed to chair
(5) - Do someone's shoelace
(6) - Manoeuver a bed
(7) - Brisk walk along corridor
(8) - Walk up stairs
(9) - Sitting down to rest

2. Overview of R-R intervals

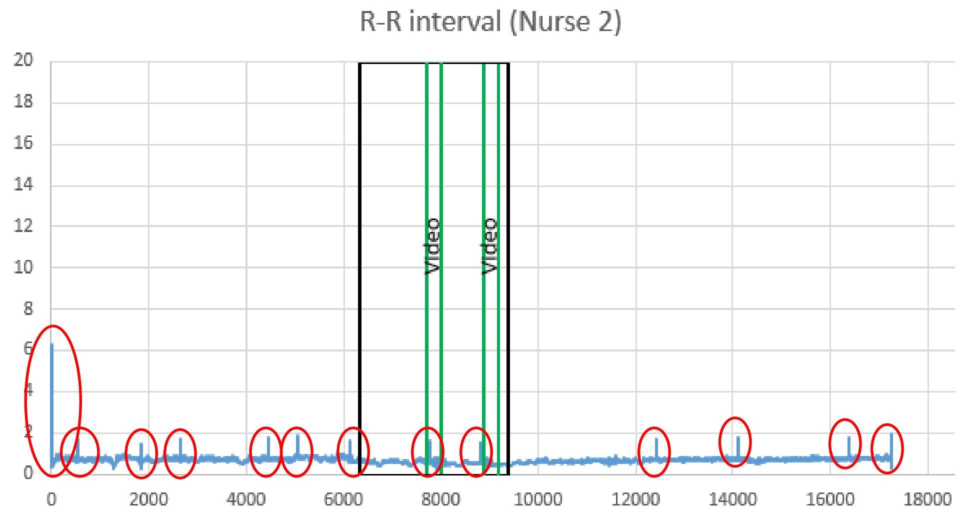


Figure 40: R-R interval of Nurse 2

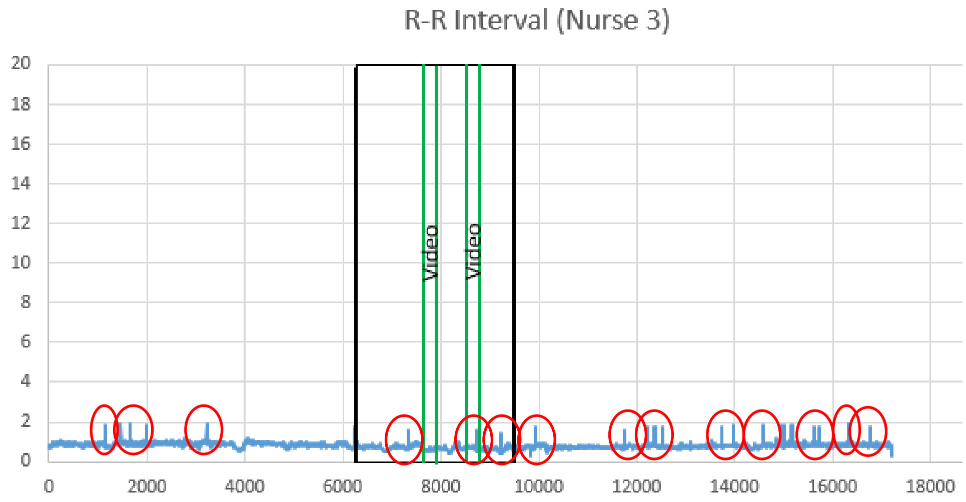


Figure 41: R-R interval of Nurse 3

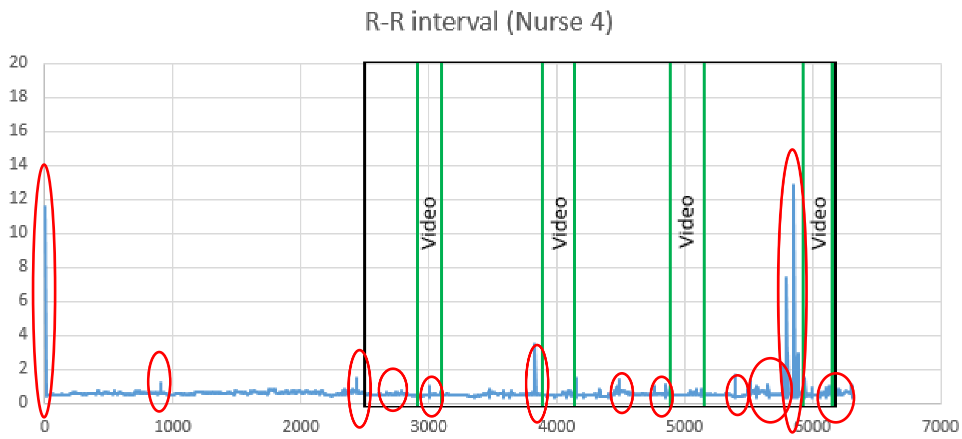


Figure 42: R-R interval of Nurse 4

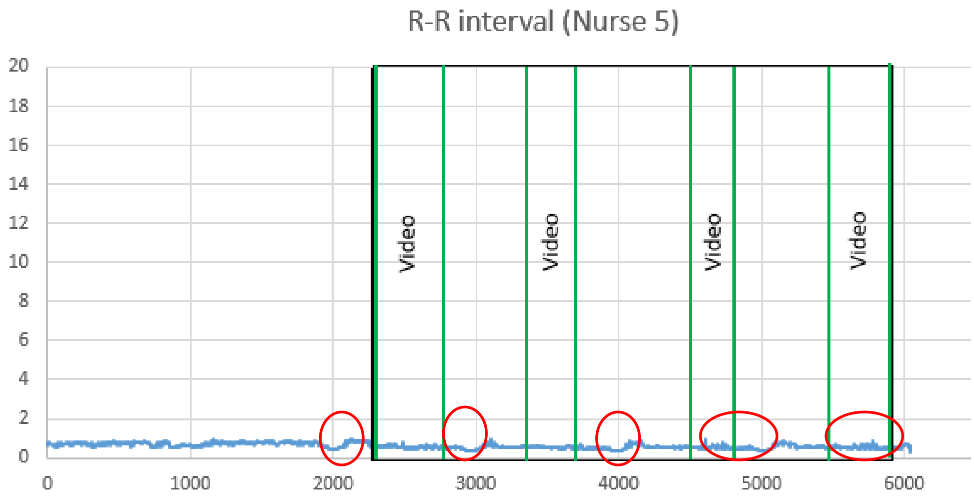


Figure 43: R-R interval of Nurse 5

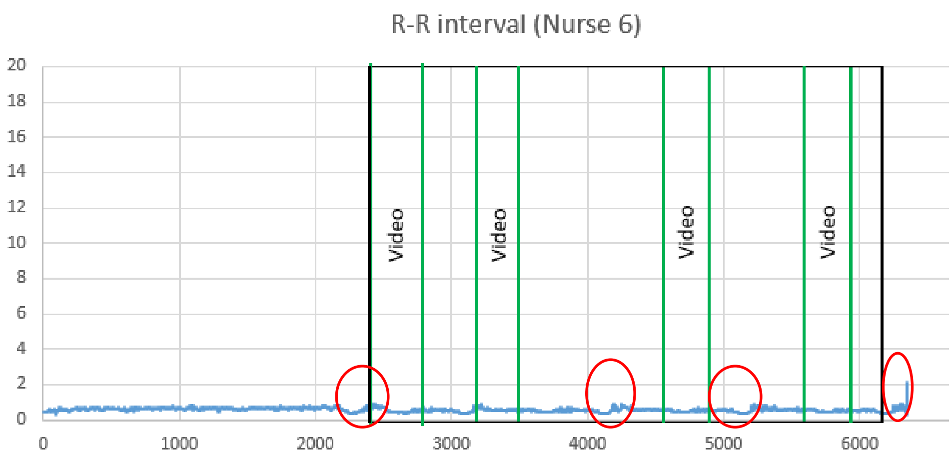


Figure 44: R-R interval of Nurse 6

3. Outlier intervals

3.1. Effect of outliers on HRV measures

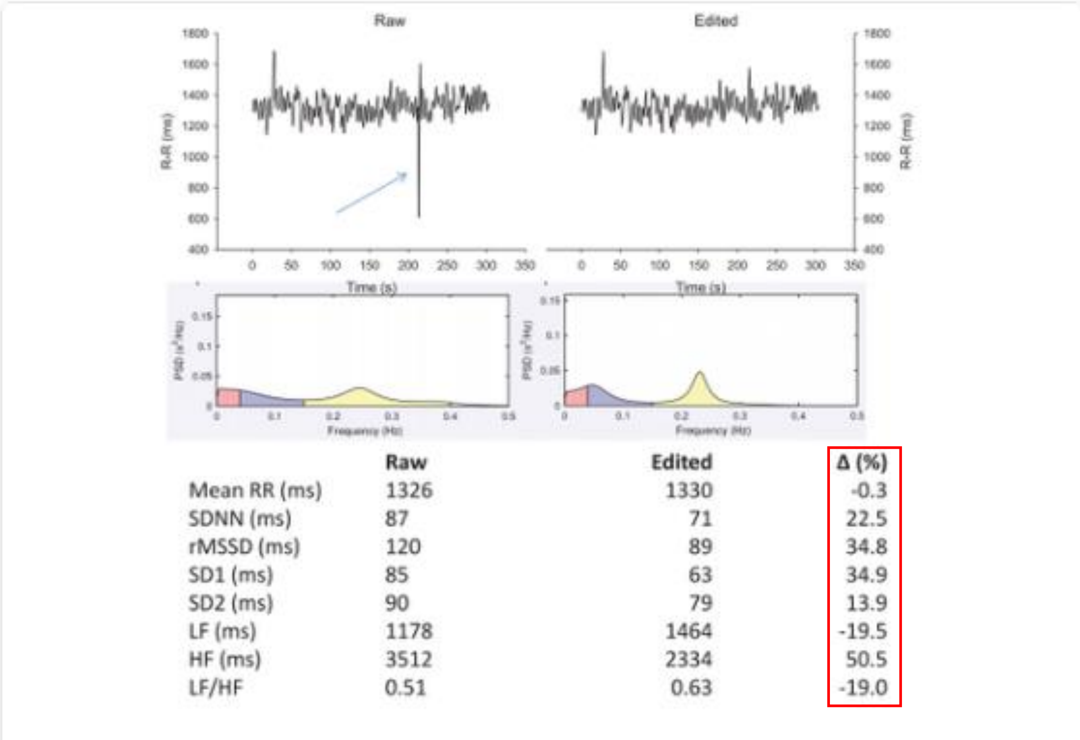


Figure 45: Comparison between raw and cleaned data (Buchheit, 2014)

3.2 Overview of outliers over all participants in the study

Recorded R-R intervals	Δ >200 ms	Δ >20%	Δ >50%
Participant 1 27498	601 (2.18%)	915 (3.33%)	222 (0.80%)
Participant 2 26545	61 (0.22%)	86 (0.32%)	54 (0.20%)
Participant 3 22384	69 (0.31%)	77 (0.34%)	60 (0.27%)
Participant 4 12183	123 (1.01%)	196 (1.61%)	96 (0.79%)
Participant 5 10640	10 (0.09%)	31 (0.29%)	5 (0.05%)
Participant 6 11560	20 (0.17%)	41 (0.35%)	7 (0.06%)

The table above displays the amount of R-R intervals of each participant:

- which differ more than 200 ms from their preceding interval.
- which differ more than 20% from their preceding interval.
- which differ more than 50% from their preceding interval.

The results show that the number of outliers recorded by the instrument is overall low. The detection of R-R intervals differentiating more than 20% from their preceding value shows the highest numbers among the filter parameters and implies a more detailed cleaning. With a look at the participants, participant 1 of the first cohort and participant 4 of the second cohort differentiate from the other participants through a higher number of outliers over all filter parameters. The higher amount of outliers existent in the data of the two participants could be caused by various reasons as mentioned in chapter 2.1.2. However, further research has to be conducted with more participants to be better able to detect patterns and draw conclusions from the results.

3.3 Application of different cleaning methods

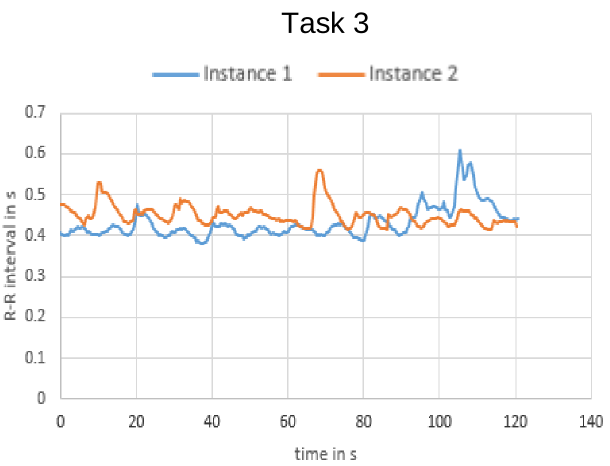


Figure 46: Nurse 1 – After removal of outliers >20%

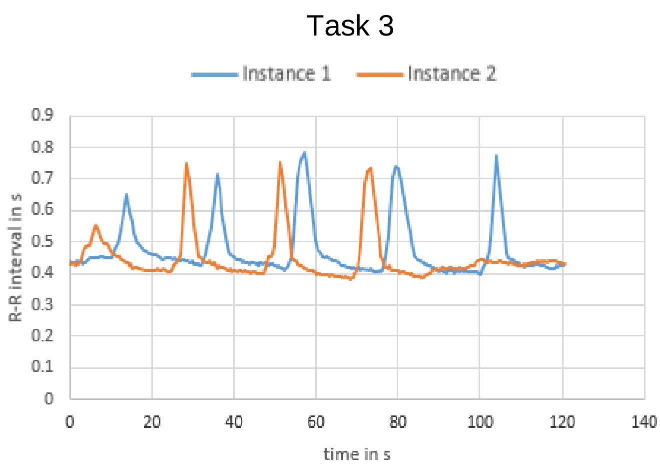


Figure 47: Nurse 2 – After removal of outliers >20%

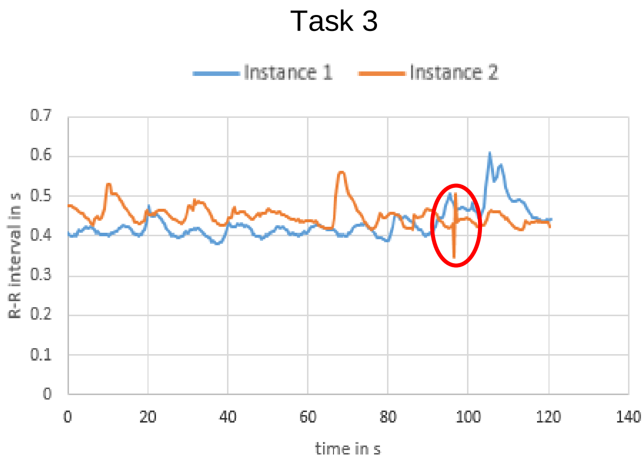


Figure 48: Nurse 1 - After removal of outliers >200ms

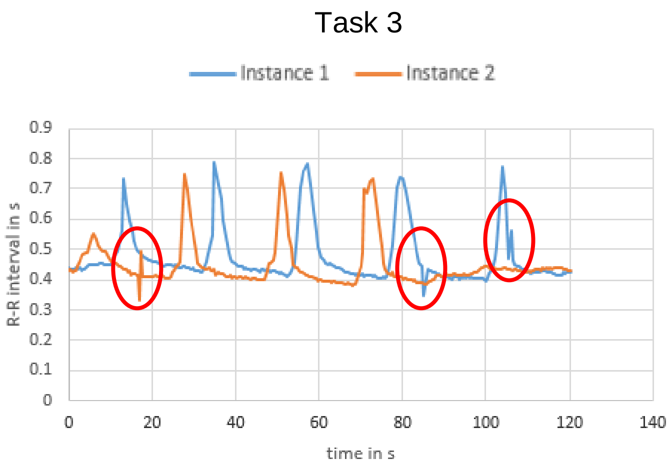


Figure 49: Nurse 2 - After removal of outliers >200ms

Figure 50: Nurse 1 – After removal of outliers >50%

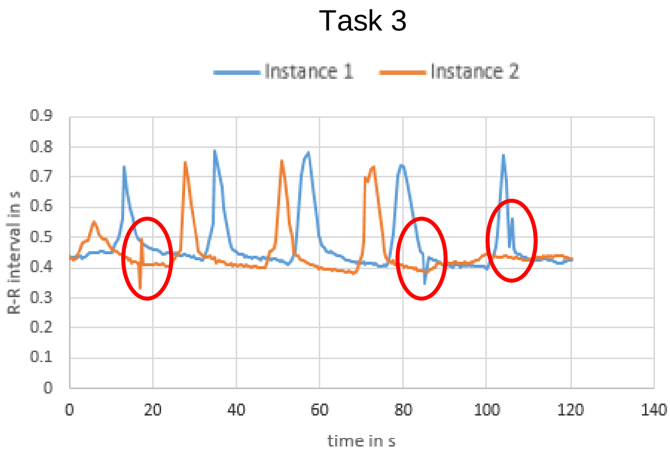


Figure 51: Nurse 2 – After removal of outliers >50%

4. Interviews and Questionnaires

4.1. Analysis of Interviews

Categories	Themes
Ease of Use & Usability	Intuitive steps to receive information about participants
	All data is joined together and easy to spot
	Aligning different data is very useful (f.e. graphs and video data)
Ability to draw conclusions from the data and contribution to further research	Tool enables to highlight exhaustion levels and patterns between participants also with regard to shift work
	Possibility to analyse tasks through a subject and a task view
	Graphs should be plotted in one diagram to be better able to compare the individual task instances
Accuracy of the Polar H7 instrument	Need to compare with a known standard or with another recording instrument
	Research needed about the reasons for the artefact to be able to get a better picture about consistency
Reliability of the Polar H7 instrument	Visualisation of the results help to compare the data over the individuals
	Polar H7 works as expected and outputs multiple datasets of the same task for a comparison
Usability of the Polar H7 instrument	Summary of the participants' answers show usability and acceptability of the collection equipment
	Allows to draw conclusions for further studies
Improvements to be made to the system	Number of cleaned outliers displayed in every view of the tool and different options to remove them from the raw data
	Tic boxes to be added to integrate the socio-demographics within the task view as well which highlight participants that match the ticked criteria
	Overlay of the task instances to make them better comparable in the subject and task view. Further overlay of the whole R-R intervals in the validation screen to compare all the recordings
	Frequency domain might be added in the further development of the tool
Benefit to further research	Real time monitoring of heart variability and stress and exhaustion levels specifically in occupations with high responsibility
	Cooperation with other institutions on a national as well as international level
	Better planning of shifts and tasks in regard to offering the highest care potential.
Python knowledge	No knowledge of Python in every interview conducted

4.2. Transcription of the interviews

Question	Interviewee	Interviewee
How do you perceive the ease of use and the usability of the system?	1	• It looks very easy to use and very intuitive in terms of clicking on a particular nurse or on a particular task that you want to receive information about. • It is very self-contained in one sort of screen so the user can see all the key vital information and can relate that just visually to the actual data in the screen with a video on top.
	2	• I think this could be really powerful and I think you explained it really clearly and we could do massive things with it. • You can see all your data together and it is simple to follow as long as you know which parameters you are looking at. I just see it is of high use especially in regard to the research we are hoping to take forward.
	3	• The tool is easy to follow and the fact that you have all the data in one place accessible. • I like the fact that you can actually link the video and the data that is coming from the devices as well.
	1	• I think the tool could be used very effectively in a number of different ways, that you either be highlighting patterns over time or identifying red flags during a shift and therefore helps people to say for example that you are not really up to the job to do this task (.) • But it's also about if you do this task and that task, what is the outcome in terms of your exhaustion level.
	2	• I think it could be used in line with current sort of topical research with regard to nurses in terms of shift patterns, levels of exhaustion,

How well does the system help you to compare the fitness level of the participants? What conclusions can be drawn from your point of view from the data through using the software?		levels of stress, reasons for outliers in the data (.) • You can see the significance of using the R-R interval really clearly now which actually when we were collecting the data also when I understood it, I could not really see it so I think the way that you broke it down is really really useful. • But I think I there are a lot of health parameters that we can look at when comparing the different R-R intervals of the nurses, how they function on a day to day basis and within their shift patterns. Really great.
	3	• The moment I can say there is high potential, think partly for the rest of the minute there are not enough participants, also I would need to explore it further how you would actually map it and show the differences between the participants. • But what I really like about this factor is that you can do the comparison between the nurses so there is a high potential in telling us lots of information about our population also in regard to their fitness.
To what extend to you think a task and a subject view can be beneficial for further comparison and hypothesis?	1	• I think there is kind of a practical explanation for that which is, it's good to be able to shift the variety of analysis, so you are looking at particular tasks (.) so that you can see that there is something about this task we want to sort of test further, but also what is highlighted is the individual within it. So if we would just have the task view, then people also might say you are only turning nursing into a task or series of tasks that are put together, whereas it is all about the individual. • So maybe it would be also good to build this up like a study over time, to compare with other recordings over time to build up a really good bank of understanding, what I suppose it is like you use the same

		task in another study, you can take that out and could compare that with another data set which has all the same tasks. • Furthermore, the graphs would further be also better comparable if they would be plotted over each other to emphasize the difference in the R-R interval of the various instances.
	2	• A task and a subject view are necessary in my eyes to see the overall picture of the study and in comparison with the demographics retrieve key facts about certain activities and individuals from the data(.) What tasks have to be investigated further, what conclusions can be drawn about the individuals, shifts, tasks and so on.
	3	• Yes, especially because you need the demographics of the subject view as well and then having a combined view would give a great value to the research.
How do you see the accuracy of the Polar H7 instrument?	1	• It seems to measure the data very accurately. It would be good if there is sort of an indication which tells us where R-R intervals are not correctly measured due to artefacts etc.
	2	• We need to know what is happening at the time of the artefact to be able to get a better picture how consistent or accurate the equipment is. It would also be nice to have two pieces of equipment to compare the findings in order to determine the accuracy of the instrument. • It would be good to see all the different plots together to quickly see all at once. Especially that you can overlay helps you accurately to see the outliers and compare with the nurses (.) Especially as well with comparing with the video.
	3	• From my experience H7 is very accurate, however there is no comparison within the tool for a verification to a known standard,

		for example a different instrument. Therefore, it is hard to say if the instrument is accurate or not.
How do you see the reliability of the used instrument for the recording of the R-R interval?	1	• Especially the visualisation of the data helps you to make a judgement about the reliability of the instrument. • For one nurse you can see there are some outliers there and you can probably explain what those were and if you can't you can adjust it. To me it gets really convincing.
	2	• Quiet well when you can overlay the graphs. So you can compare one nurse to the other and see what is happening with the R-R interval, if there are outliers, why they appear especially with regard to the video. You further can adjust the outliers so out of my point of view (.) the tool enables the user to judge the reliability of the polar device.
	3	• The H7 can be seen as reliable and cheap for monitoring HRV. It can be seen as reliable, because it works as expected and there are multiple datasets of the same task for a comparison. • The reliability can also further be determined in regard to the recordings of other participants in the study (.) Also in the validation view a general plot with all the data would help to better be able to compare the recordings.
To what extend does the system enable you to make judgements about the usability of the collection equipment (Polar H7 instrument)?	1	• Yes, especially the attempts to take for the instrument. In the study it seemed to be fine and then you are looking and it is actually talking three attempts in average and that means do we have to show them how to put it on.
	2	• It is a good complement to the qualitative collective data that has to be collected with regards to the usability and acceptability, so you are complementing that data

		really well with what you are presenting.
	3	• I think you can judge the usability with looking at the summary of their answers in the screen. This is especially complementary to the qualitative work done in this project.
What would you improve in regard to the system and why?	1	• In my eyes it would be good to plot the graphs in one plot to be able to better compare them especially when the variations between the instances are rather low (.) Also the number of cleaned outliers should be displayed. Especially when reporting it you could say, across the 6 nurses there were on average x outliers for each, so that would be definitely beneficial as well. • Furthermore, it would be beneficial to tag participants based on key socio-demographic characteristic for example in the task view to tic boxes for smoker, fitness activity (.) and nurses would be tagged when there is a fit.
	2	• A relationship between the demographics and the graphs would be beneficial as well as an overlay of the graphs (.) • The frequency-domain could be also implemented at some point due to more accurate information in terms of numeric stress recording. The plots could also be changed to frequency domain within the further development of the tool.
	3	• Overlaying all the data together, perhaps having a function at the bottom of the validation view as an extra one (.) So you can keep the individuals and have an extra plot over all the individuals. This then doesn't have to be necessarily linked to the video because you could see in that plot, oh there is a problem with

		Nurse 3 for example and go back to the individual view. • Also the graphs in the task and the subject view could be overlaid to better compare the R-R intervals.
To what extend does the system benefit to further research?	1	• For example, this could also be used at some point in real time especially with regard to the task view. Just image areas with high responsibility you have a surgery team or in air traffic control and you could monitor their stress, exhaustion level, you could switch them.
	2	• I think this tool is applicable for various groups of people. Especially also in regard to cooperations not just within the UK but also globally, it has much potential.
	3	• So you could see for the nurses, this nurse was doing everything which was really not good for them or had a higher exhaustion(.) so you might rethink their shifts and tasks in regard to offering the highest care potential.
How do you define your knowledge of Python?	1	• Don’t know anything about it.
	2	• Never heard of it to be honest.
	3	• How I define my knowledge of Python? I don’t have any knowledge about it.
What is your opinion about the requirements of the system? What functionalities would you add or delete?	1	• The tic boxes would be a nice additional function as earlier discussed (.) Also to know the cleaned number of outliers would be another thing (.) the user can remove outliers in the data within every view to have a comparison between cleaned and uncleaned data.
	2	• No not really, I think so far we discussed everything important and the functionalities look good.
	3	• An overlay of the R-R graphs in the validation view would be good to enable an easy comparison (.) Also graphs for the time-domain measures should be implemented

		to have additional visualisations to compare the data. Another thing would be the outliers (.) The user should have more flexibility within the cleaning options (.) from less aggressive to more aggressive cleaning.

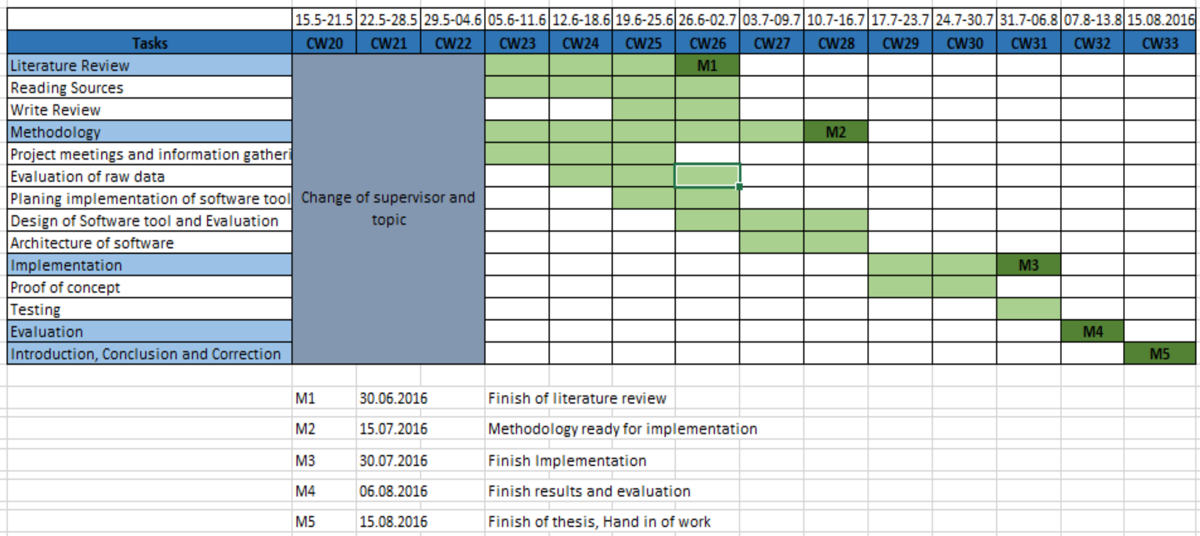
4.3. Questionnaire answers

Statement	Interviewee	Questionnaire rating on a scale from 1 (Strongly agree) to 7 (Strongly Disagree)
Thesystem is easy to learn and has a high usability	1	1
	2	2
	3	1
The system helps to compare the HRV of the participants to draw conclusion about the data	1	1
	2	1
	3	2
A task and a subject view is beneficial in this system	1	1
	2	1
	3	1
The system helps to determine the accuracy of the used instrument (Polar H7)	1	3
	2	4
	3	7
The system helps to determine the reliability of the used instrument (Polar H7)	1	2
	2	1
	3	1
The system helps to make judgements about the usability of the used instrument in the study from a participant's perspective	1	1
	2	1
	3	1
The system can be seen as a foundation for further research	1	2
	2	1
	3	2
I have knowledge about the Python programming language	1	7
	2	7
	3	7
The requirements for the system are satisfying	1	2
	2	2

	3	3

The red fields show the answers given in the questionnaire, which don't match the answers given in the interviews by the participant.

5. Project Management



6. Research proposal

6.1 Student details

Last (family) name	Schwab
First name	Christian
Napier matriculation number	40217437

6.2 Details of your programme of study

MSc Programme title	Information Systems Development
Year that you started your diploma modules	2015
Month that you started your diploma modules	September
Mode of study of diploma modules	Full-time
Date that you will complete your diploma modules at Napier	29 – April – 2016

6.3 Project outline details

Please suggest a title for your proposed project. If you have worked with a supervisor on this proposal, please provide the name. NB you are strongly advised to work with a member of staff when putting your proposal together.

Title of the proposed project	**An evaluation of Python tools for analysing wellbeing data**
Is your project appropriate to your programme of study?	Yes
Name of supervisor	Alistair Lawson
I do not have a member of staff lined up to supervise my work	No

6.4 Brief description of the research area – background

Please provide background information on the *broad research area* of your project in the box below. You should write in narrative (not bullet points). The academic/theoretical basis of your description of the research area should be evident through the use of references. Your description should be between half and one page in length.

This Masters dissertation project aims to develop and evaluate a Python software tool for analysing, integrating, cleaning, aligning and visualizing wellness data relating to the performance of nurses under various working scenarios. The project will use data that was previously collected using the Polar Loop, a wristband tracking the heart rate (R-R interval and BPM), speed, distance, acceleration and cadence. Other data to be integrated includes Video and Questionnaire data.

Nursing is a challenging profession, which includes not only high professional performance prospects, but also a high extent of public performance expectation beside its long varying working hours and high level of stress that are often associated in this profession (Schluter et al., 2011; Videman et al., 2005). The resulting mental stress and physical exposure in this job are therefore substantial for health issues and physical wear over time (McNeely, 2005).

Information technology has the potential to help improve nursing practices and efficient care (Huntington et al. 2009; Gelsema et al., 2006). This includes the use of technology to analyse data to detect or minimise stress levels and improve the physical state of nurses is rather muted (Yuan et al., 2009; Edwards & Burnard 2003). This project focuses on Python, an open source, multi-purpose scripting technology (Grus, 2015; Mehlhase, 2014) which can be used to develop tools to integrate, clean, align, analyze and visualize the different retrieved types of data: video data (recording of different activities), numerical data (analyses of Polar Loop) and questionnaire data (test persons).

Generally, data can be divided in structured data, unstructured data and semi-structured data (Assuncao et al., 2015). Structured data as the retrieved numerical data can be described as data allocated to fields and therefore can be easily processed. Nevertheless, most of the generated data is semi-structured or unstructured as the questionnaires or video material in this project (Baars & Kemper, 2008; Assuncao et al., 2015).

6.5 Project outline for the work that you propose to complete

Please complete the project outline in the box below. You should use the emboldened text as a framework. Your project outline should be between half and one page in length.

The idea for this research arose from:

Meetings with my supervisor.

The aims of the project are as follows:

- This Masters dissertation project aims to develop and evaluate a Python software tool for integrating, cleaning, aligning, analyzing, and visualizing wellness data relating to the individual performance of nurses under various working scenarios.

The main research questions that this work will address include:

- What questions might be asked of the data?
- Does the equipment (e.g. the Polar loop) ensure a reliable data recording?
- To what extent does the software tool enable research to draw conclusions about the data and provide a foundation for further investigations in this area?

Other deliverables of the project will be:

- Test results.
- Tool in python for the analysis of unstructured and structured data, providing a foundation for further research in this area.

This work will require the use of specialist software: Python.

6.6 References

Please supply details of all the material that you have referenced in sections 4 and 5 above. You should include at least **three references**, and these should be to high quality sources such as refereed journal and conference papers, standards or white papers. http://www.soc.napier.ac.uk/~cs104/mscdiss/moodlemirror/d2/2005_hall_referencing.pdf

Assuncao, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A., Buyya, R. (2015). Big data computing and clouds: trends and future directions, *Journal of Parallel and Distributed Computing*, 79, 3-15.

Baars, H. & Kemper, H-G. (2008). Management Support with Structured and Unstructured Data An Integrated Business Intelligence Framework, *Information Systems Management, 25(2),* 132-148.

Edwards, D. & Burnard, P. (2003) A systematic review of stress and stress management interventions for mental health nurses. *Journal of Advanced Nursing, 42* (2), 169–200.

Grus, J. (2015). *Data Science from Scratch: First Principles with Python*. Beijing: O'Reilly.

Gelsema, T., Van der Doef, M., Maes, S., Janssen, M., Akerboom, S. & Verhoeven, C. (2006). A longitudinal study of job stress in the nursing profession: causes and consequences. *Journal of Nursing Management*, *14*, 289–299.

Huntington, A., Gilmour, J., Schluter, P., Tuckett, A., Bogossian, F. & Turner, C. (2009). The Internet as a research site: establishment of a web-based longitudinal study of the nursing and midwifery workforce in three countries. *Journal of Advanced Nursing*, 65 (6), 1309–1317.

McNeely, E. (2005). The consequences of job stress for nurses' health: time for a check-up. *Nursing Outlook, 53*(6), 291–299.

Mehlhase, A. (2014). A Python framework to create and simulate models with variable structure in common simulation environments, *Mathematical and Computer Modelling of Dynamical Systems, 20*(6), 566-583.

Schluter, P.J., Turner, C., Huntington A.D., Bain, C.J. & McClure, R.J. (2011) Work/life balance and health: the Nurses and Midwives e-cohort study. *International Nursing Review* 58, 28–36.

Videman, T., Ojajarvi, A., Riihimaki, H., Troup, J., (2005). Low back pain among nurses: a follow-up beginning at entry to the nursing school, *Spine, 30*(20), 2334-2341. Yuan, S.C, Chou, M.C., Hwu, L.J, Chang, Y.O., Hsu, W.H., Kuo, H.W. (2009). An intervention program to promote health related physical fitness in nurses. *Journal of Clinical Nursing*, *18*, 1404-1411.

6.7 Ethics

If your research involves other people, privacy or controversial research there may be ethical issues to consider (please see the information on the module website). If the answer below is YES then you need to complete a research Ethics and Governance Approval form (available on the website: http://www.ethics.napier.ac.uk).

Does this project have any ethical or governance issues related to working with, studying or observing other people? (YES/NO)	**YES** The project will use anonymised data that was collected following ENU ethical procedures by School of Nursing Midwifery and Social Care. This has been through the ethical approval process

Supervision timescale

Please indicate the mode of supervision that you are anticipating. If you expect to be away from the university during the supervision period and may need remote supervision, please indicate.

Weekly meetings over 2 trimester	No
Meetings every other week over 3 trimesters	Yes
Other	-