Heart and Respiration function relate d to Age and Gender

By

IPSITA GHOSH

Bachelors of Technology in

Information Technology,

Future Institute of Engineering and Management

West Bengal, India

2016

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
In partial fulfillment of
the requirement for
the Degree of
MASTER OF SCIENCE
July, 2019

Heart and Respiration function related to Age and Gender

Thesis Approved:

Dr. Johnson P. Thomas

Thesis Adviser

Dr. K. M. George

Dr. Bruce Benjamin

ACKNOWLEDGEMENTS

The Master's degree from the Computer Science department, Oklahoma State University has given me immense experience and knowledge in my field of interest. I would like to thank my thesis advisor, Dr. Johnson Thomas for his continuous assistance and encouragement when I was doing my research.

I would like to express my gratitude to my committee members, Dr. Bruce Benjamin and Dr. K. M. George for their guidance and support. Dr. Benjamin in particular has been a great source of ideas and encouragement.

Finally, I wish to express my profound gratitude to my family and friends who supported me throughout my years of study.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: IPSITA GHOSH

Date of Degree: JULY, 2019

Title of Study: HEART AND RESPIRATION FUNCTION RELATED TO AGE AND GENDER

Major Field: COMPUTER SCIENCE

Abstract

In the human body, all the organs are interconnected and interdependent. In this work we consider the heart and the lungs. Changes in the functionality of one can affect the other. The heart controls blood flow and the lungs controls respiration.

In the thesis we investigate if heart rate variability (HRV) and respiratory rate variability (RSV) to can give a more accurate prognosis of the state of the body. Previous research has focused on HRV. As a first step we see if both HRV and RSV can predict gender and age of a person more accurately than HRV or RSV alone. In this work we looked at healthy people.

We analyzed the data of healthy people from the PhysioNet database [14]. Our results show that Gender influences Heart Rate and Breathing Rate. Consider both together rather than individually give a better prediction of gender. The only feature that is related to age is the standard deviation of respiration rate.

TABLE OF CONTENTS

List of Figures

Page

List of Tables

Page

CHAPTER I

INTRODUCTION

1 Introduction

The organs in the body are all interconnected. A change in one can affect the other. Monitoring organs is therefore essential for well-being. Abnormality in any of the organ functions may cause problems in different organs. The heart and lungs are very powerful organs of the human body. In this research we look at how the condition of the heart and lungs can help in predicting the difference between male and female of different age group health wise.

Information about heart and lungs can provide how the human body is functioning. The heart gives the information about the cardio-vascular system and lungs give information about the respiratory system. These two are interconnected and changes in one may affect the other. The long term goal of this research is to predict the probability of cardiovascular or lung diseases based on Electro-Cardiogram (ECG) and Respiration (inspiration and expiration) and other factors such as age and gender. The immediate goal is to see if these readings can be used to determine the gender and age of a person. Gender and age play an important role is disease manifestations. Men for example are prone to heart attacks at an earlier age when compared to women. Hence, distinguishing gender or age based on the data will lead to more accurate analysis for diseases. A key question that this research seeks to answer is if analysis based on both heart and lung data will deliver more accurate

classification of gender or age or if analysis of only one of heart or lung data will provide better classification.

Currently, there are many ways to monitor the health of heart and lungs. Our long term goal is that predict heart or lungs health condition based on age and gender and how it changes depending upon health issues. For our research a dataset is obtained from the databank of Physionet [ref]. The data set is called Fantasia data. It has 40 subjects of 20 males and 20 females, 20 young and 20 old people. The subjects are all healthy as far as we are aware. The dataset has ECG and Respiration information of these 40 subjects. The data is collected for every 4 milliseconds for 2 hours while these 40 people were watching a Disney movie called 'Fantasia'. This data is executed individually one by one through a pre-built software called PUKA to collect detailed information about the heart and lungs of these 40 people.

ECG gives all heart and heart-rate related information and respiration gave all respiratory and respiratory-rate related information. Our goal is to find out whether respiratory rate variability (RSV) combined with the heart rate variability (HRV) improves the gender or age classification of patients.

In chapter II we describe the essential functioning of the cardiovascular system. In chapter III we outline the essential functioning of the respiratory system. Chapter IV presents a literature review on work done on heart-rate variability and respiratory-rate variability separately. In chapter V we state our hypothesis and the problem statement. We propose the methods and model to solve the

problem and to achieve our goal. Results are also discussed in this chapter. In chapter VI we state our findings and conclusion.

CHAPTER II

CARDIOVASCULAR SYSTEM

2 Cardiovascular System

The **cardiovascular system** consists of the **heart**, blood vessels, and blood. These three **main functions of the system are**: Transport of nutrients, oxygen, and hormones to cells throughout the body and removal of metabolic wastes (carbon dioxide, nitrogenous wastes).

2.1 HEART RATE

Heart rate is the speed of the heartbeat measured by the number of contractions (beats) of the heart per minute (bpm). The heart rate can vary according to the body's physical needs, including the need to absorb oxygen and excrete carbon dioxide.

The American Heart Association states the normal resting adult human heart rate is 60–100 bpm. Tachycardia is a fast heart rate, defined as above 100 bpm at rest. Bradycardia is a slow heart rate, defined as below 60 bpm at rest. During sleep a slow heartbeat with rates around 40–50 bpm is common and is considered normal. When the heart is not beating in a regular pattern, this is referred to as an arrhythmia. Abnormalities of heart rate sometimes indicate disease.

2.2 THE HEART RATE DEPENDS UPON A FEW FACTORS:

a) **Air temperature:** When temperatures (and the humidity) soar, the heart pumps a little more blood, so the pulse rate may increase, but usually no more than five to 10 beats a minute.

b) **Body position:** Whether resting, sitting or standing, the pulse is usually the same. Sometimes as you stand for the first 15 to 20 seconds, the pulse may go up a little bit, but after a couple of minutes it should settle down.

c) **Emotions:** Stress, anxious or "extraordinarily happy or sad" emotions can raise pulse.

d) **Body size:** Body size usually doesn't change pulse. Highly obese persons may see a higher resting pulse than normal, but usually not more than 100.

e) **Medication use:** Medicines that block adrenaline (beta blockers) tend to slow pulse, while too much thyroid medication or too high of a dosage will raise it.

f) **Autonomic nervous system:** An increase in parasympathetic nervous system activity decreases heart rate and an increase in sympathetic nervous system activity increases heart rate. Changes in the autonomic nervous system are initiated by receptors located in the body to match heart rate with changes in the demands of the cardiovascular system.

2.3. *Electrocardiogram* (ECG) is a medical test that detects cardiac (heart) abnormalities by measuring the electrical activity generated by the heart as it contracts. The machine that records the patient's ECG is called an electrocardiograph. The echocardiogram and EKG are very important and are commonly used heart tests that are instrumental in diagnosing various forms of heart disease. The echocardiogram is an ultrasound of the heart that provides moving pictures and provides information on the structure and function of the heart. The EKG is a heart tracing that

mainly provides information on the rhythm of the heart. Both tests are often used in conjunction and are complimentary to each other.

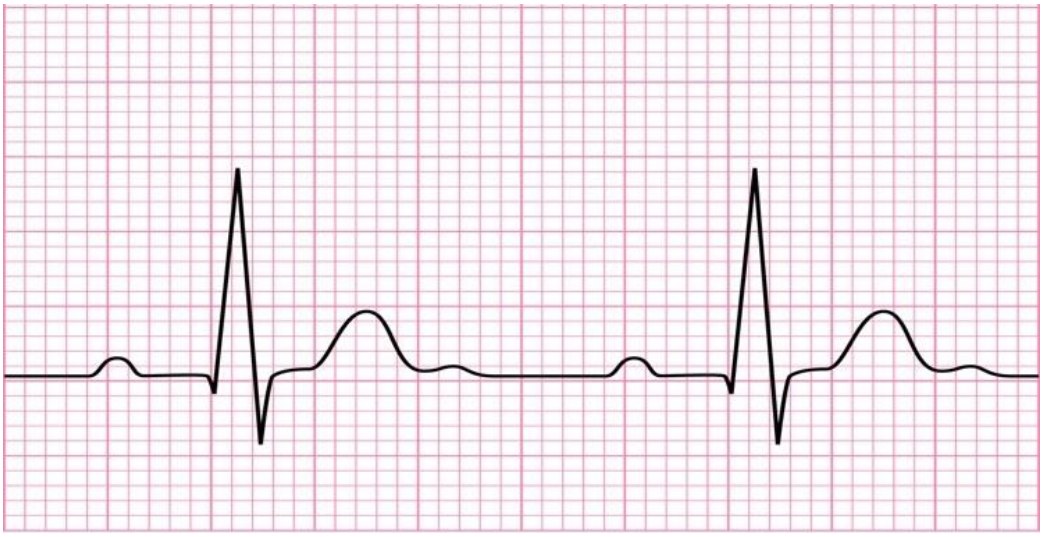

Figure 1: ECG Record

2.4. Heart Rate Variability is a measure which indicates the variation in heartbeats within a specific timeframe. The unit of measurement is milliseconds (ms).

2.5. The technical term for calculation of heart rate is *R-R interval*. It is the instantaneous time between two R peaks in ECG or QRS complexes. In normal cases, the time between two r waves is one second. In general, the normal interval time is 0.4 to 0.44 seconds. Women have longer interval than men.

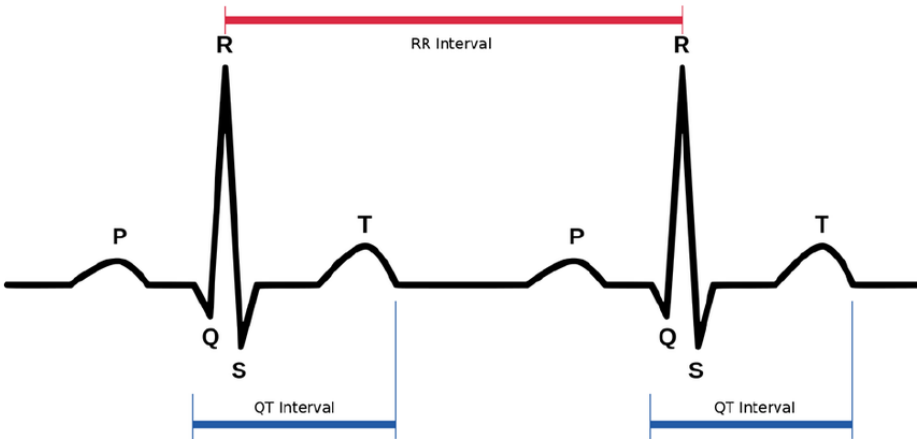

Figure 2: R-R Interval

Chapter III

RESPIRATORY SYSTEM

3 Respiratory System:

The respiratory system allows a person to breath and exchange oxygen and carbon dioxide throughout the body. The human respiratory system includes nose (nasal passages), larynx, trachea, bronchial tubes, lungs and diaphragm. The respiratory system is also called as the **Gas Exchange System.** Through this system, the body gets rid of carbon dioxide and takes in oxygen.

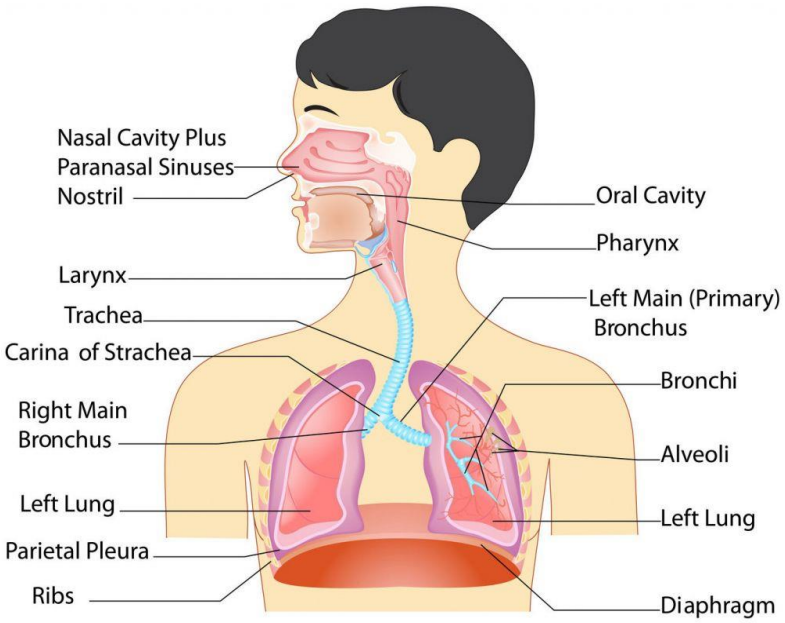

Figure 3: Respiratory System

3.1 RESPIRATORY RATE

This is the number of breaths per minute or, more formally, the number of movements indicative of inspiration and expiration per unit time. In practice, the respiratory rate is usually determined by counting the number of times the chest rises or falls per minute. The normal respiration rate for an adult at rest is 12 to 20 breaths per minute. A respiration rate under 12 or over 25 breaths per minute while resting is considered abnormal. A normal breathing rate for an adult at rest is 8 to 16 breaths per minute. For an infant, a normal rate is up to 44 breaths per minute. The term hyperventilation is usually used if you are taking rapid, deep breaths. This can be due to lung disease or because of anxiety or panic. The process of breathing (respiration) is divided into two distinct phases, inspiration (inhalation) and expiration (exhalation).

3.2. Inspiration or inhalation is the process of obtaining atmospheric oxygen. In this process the diaphragm drops a little and the muscles in the rib cage move up a little. This causes the volume of the thoracic cavity to increase. Since the volume of thoracic cavity increases the pressure inside the lungs drops below the atmospheric pressure. The difference in pressure forces the atmospheric air to rush inside the lungs through nasal cavity and wind pipe.

3.3. Expiration or exhalation is the process of releasing carbon dioxide to the atmosphere. In this process the diaphragm relaxes (or moves up) and the muscles in rib cage moves down a little. This causes the volume of the thoracic cavity to decrease. Since volume of thoracic cavity decreases the pressure inside the lungs rises above atmospheric pressure. The difference in pressure forces the carbon dioxide to rush out of the lungs through nasal cavity and wind pipe.

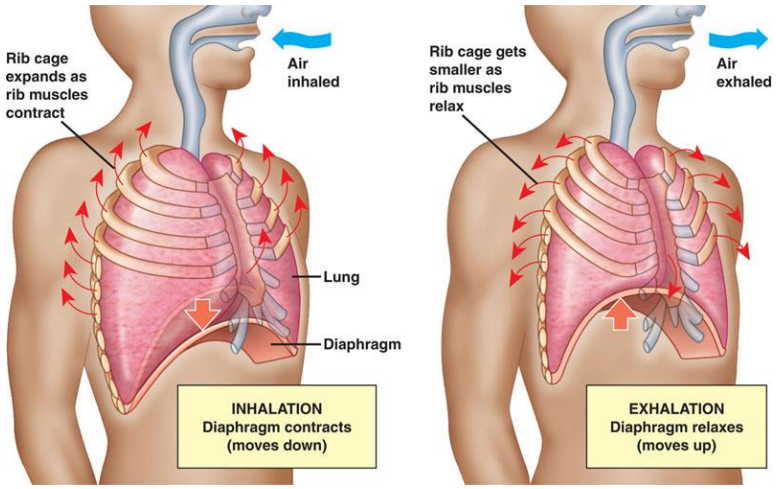

Figure 4: Inhalation and Exhalation

Chapter IV

REVIEW OF LITERATURE

4 Related work

Heart rate variability (HRV) is associated with both heart rate and respiratory rate. Changes in any one of these two or both may cause a change in HRV. Heart rate variability (HRV) is a non-invasive method to measure cardiac autonomic function. HRV is the effect of the changes in the measurement of heart rate and respiratory rate. In the stable condition where no external condition is affecting heart rate, heart rate is the main factor to determine HRV. In the stable condition respiratory rate is not a factor in determining HRV. A small change in heart rate may affect HRV.

A reduced HRV can cause cardiovascular disease. Cardiovascular disease is one of the leading causes of death. Older people are more likely to be afflicted with cardiovascular disease. HRV decreases with increasing age. A reduced HRV is associated with a higher risk of coronary heart diseases which leads to heart failure. HRV is influenced by heart rate and respiratory rate. Multiple regression analysis says that only heart rate alteration is proved to be an independent determinant of HRV [1]. Every change in heart rate by 1 bpm changes the HRV values by 16.5% on average.

Low HRV is considered an independent marker of mortality risk. Studies say that age related decline in HRV may limit its predictive value, especially in elderly people. HRV decreases linearly with aging. HRV in young female is lower than in young male. The effect of gender decreases when age is more than 30 and slowly disappears when age is more than 50. But heart rate (HR) decreases much more slowly with aging. HR is faster in female than in male at the age less than 50. Gender differences disappears after that. HRV of healthy people declines with aging. Decreasing HRV of healthy people is associated with increased risk of mortality with growing age (age > 65). Gender differences in HRV are age and measure dependent. Age and gender also affect heart rate. [1]

At young age especially in children the standard deviation of normal to normal (SDNN) R-R interval and the standard deviation of the average normal to normal interval (SDANN) is significantly higher in male than female. With increasing age there is a progressive and significant decrease of HR and increase of SDANN. HRV changes depending upon age and gender. Up to the age of 10, gender is unrelated to HRV. [2]

Though there has been a decline in rates of mortality due to heart disease during the past two decades, cardiovascular disease remains the most frequent single cause of death among persons over 65 years of age. It also accounts for a major and growing proportion of hospitalizations and health care costs and is a leading cause of morbidity among older people. Considering change in age, may influence the diagnosis and the treatment of cardiac diseases that is the treatment may vary between older and younger patients. [3]

There were many computerized methods designed to investigate the effect of age upon heart rate variability at rest and in response to a single deep breath. A computerized method of measurement of R-R interval variation was used to study heart rate responses in healthy subjects aged 18-85 years. Heart rate variation during each procedure showed a skewed distribution and a statistically significant negative correlation with age. Normal ranges (90% and 95% confidence limits) for subjects aged 20-75 years were calculated for heart rate difference (max-min) and ratio (max/min) and standard deviation (SD). [4]

There are researches which shows the relationship between input and output of a cardiovascular system depending upon age with no cardiovascular disease. Heart rate and end-diastolic aortic pressure remained unchanged with age, whereas aortic systolic, mean and pulse pressures and aortic radius increased. In subjects younger than 30 years, early systolic pressure usually exceeded late systolic pressure; in subjects older than 50 years, late systolic pressure usually exceeded early systolic pressure. In the study 55% of the subject is between 30 to 50 years and the late systolic pressure were equal for them. So age generally effects the cardiovascular system at older age. [5]

The HRV is relative to age, gender, heart rate, body mass index etc. without any prior knowledge of heart diseases. There was an inverse correlation of HRV with heart rate ($p < 0.001$). HRV indexes decreased with increasing age, also differed by gender, and were higher in patients with higher functional capacity. No correlation was noted between HRV and body mass index. [6]

Impairments in HRV have been proposed as independent risk factor for increased cardiac mortality and morbidity. In studies [12] it is assessed that the effect of age and gender on

autonomic regulation of heart in healthy volunteers. None of the subjects had medical illness such as diabetes, hypertension, thyroid disorders, cardiac disorders etc. HRV is negatively corelated with age. Multiple regression analysis is done to control for effect of age and heart rate while comparing males and females and significant reduction came out that sympathetic tone is lower in females. There is an overall reduction in autonomic control of heart with increase in the age. Females showed greater vagal tone than male. This differential autonomic tone indicate age, gender related predisposition to cardiovascular disease. [7]

Studies of effects of aging on the respiratory system may be difficult to interpret for several reasons. Chronic exposure to environmental pollutants, repeated pulmonary infections, smoking, and differences in life- style, working conditions, and socioeconomic factors may cause alterations in the respiratory system that are not easy to distinguish from changes due to aging alone. [8]

Respiratory control is influenced by age and gender. Specifically, respiratory long-term-facilitation following intermittent hypoxia decreases with age in male, but increases in female. This experiment is done on rats. It's difficult to give information related to human age and its effect on respiratory control. [9]

The pulmonary function is distinguishable depending upon gender with the presence of some other external criteria. Studies say during heavy exercise, women demonstrate greater expiratory flow limitation, an increased work of breathing compared to men. [10] Gender affects the incidence, susceptibility and severity of several lung diseases. Gender also influences lung development and physiology. [11]

Physiological ageing of the lung is associated with dilatation of alveoli, enlargement of airspaces, decrease in exchange surface area and loss of supporting tissue for peripheral airways. [15] The changes result in decreased elasticity of lungs. Respiratory muscle strength also decreases with ageing, and is strongly correlated with nutritional status and cardiac index. Expiratory flow rates decrease with a characteristic alteration in the flow-volume curve suggesting small airway disease. Respiratory function results indicate that some measures are different for women than for men. [16] Respiratory results indicate that tracheal pressure, percent rib cage contribution, lung volume, and rib cage volume initiations are higher, and lung and rib cage volume excursions are larger when higher vocal intensity levels are produced.

CHAPTER V

PROBLEM STATEMENT AND TECHNIQUES USED

5 Problem Specification

There are many factors that differentiate the health condition of a human cardiovascular system and respiratory system. It varies age to age and between genders. Hence we measure heart and lungs condition separately for age and gender.

The primary goal of our study is to analyze and predict the differences based on cardiac and respiratory information of old and young, male and female.

5.1 PROPOSED SOLUTION

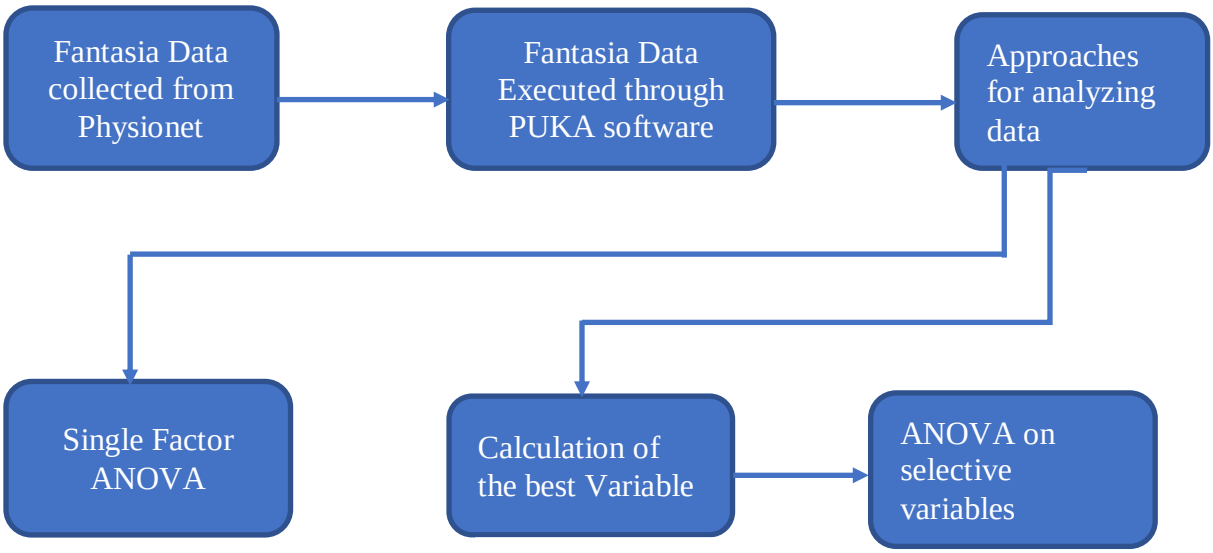

Figure 5: Flow Chart

5.2 DATA

Physionet is a forum for the exchange of data and software among researchers. PhysioNet offers free access via the web to large collections of recorded physiologic signals and related open-source software. The PhysioNet web site is a public service of the PhysioNet Research Resource for Complex Physiologic Signals, funded by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health.

Physiobank is a large and growing archive of well-characterized digital recordings of physiologic signals, time series, and related data for use by the biomedical research community. PhysioBank currently includes more than 60 collections of cardiopulmonary, neural, and other biomedical signals from healthy subjects and patients with a variety of conditions with major public health implications, including sudden cardiac death, congestive heart failure, epilepsy, gait disorders, sleep apnea, and aging. These collections include data from a wide range of studies, as developed and contributed by members of the research community.

5.2.1 Fantasia Database

Twenty young (21 - 34 years old) and twenty elderlies (68 - 85 years old) rigorously-screened healthy subjects underwent 120 minutes of continuous supine resting while continuous electrocardiographic (ECG), and respiration signals were collected; in half of each group, the recordings also include an uncalibrated continuous non-invasive blood pressure signal. Each subgroup of subjects includes equal numbers of men and women [13]. All subjects remained in a resting state in sinus rhythm while watching the movie *Fantasia* (Disney, 1940) to help maintain

wakefulness. The continuous ECG, respiration, and (where available) blood pressure signals were digitized at 250 Hz. Each heartbeat was annotated using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection.

5.2.2 PUKA Software

Puka is a Hawaiian word meaning *appear*, *emerge*, or *an opening*, and is also the name of a small round shell with a center hole found on some Hawaiian beaches. It was selected to reflect the exploratory nature of the experiments which generate the data to be analyzed with this program as well as the new knowledge that may be gained. Puka calculates descriptive statistics such as heart rate variability (HRV), peak-valley respiratory sinus arrhythmia (RSA), and respiratory variables from EKG and strain gauge respiration data.

Puka, written in Java, uses MATLAB for signal processing and statistical calculations, and the WFDB Software Package and ecgpuwave for EKG peak detection. Puka can also read and write data from a MySQL database. Puka was developed and tested using Cygwin under MS-Windows, but Puka should be usable on other platforms supported by MATLAB as well.

Puka incorporates a new method of identifying the breaths and pauses in strain gauge belt recordings. This technique locates the points of maximum inspiration and expiration for each breath as well as post-inspiratory and post-expiratory pauses. Puka correctly locates normal R waves in EKG signals and breaths in strain gauge belt recordings, in tests using artificial EKG data, paced respiration recordings from healthy young subjects, and recordings from neurological patients. [12]

We used Puka to extract Heart Rate variability (HRV) and Respiratory Rate Variability (RSV) from the Fantasia data. Puka calculated Mean of Inspiration and Expiration Time, Standard deviation for Inspiration and Expiration Time, Number of Breath in particular time frame, Breathing Rate, Shortest breath and Longest breath in a particular time frame for RSV. It also calculated R-R Interval Mean and R-R Standard deviation, Heart Rate, Shortest Beat, Longest Beat in a particular time frame for HRV.

PUKA is the software that is used in our work. PUKA works with the support of Java, MatLab, Cygwin, WFDB, JMatLink, MySQL etc. This is an old software to work with. All the other tools and platforms were updated except the JMatLink. JMatlink connects MatLab and Java. The general idea is to use MATLAB as a pivot in a design and measuring and data processsing environment in connection with real-time hardware. The task was to connect java librabies with MatLab engine. JMatLink libraries are built to connect java with MatLab. JMatLink [18] is a very old library and has not updated since 2003. This created a lot of problems in running Puka, we found a github code link for jmatlink. It worked as an adopter to connect java with MatLab so that Puka can work. Before this adopter was installed Puka was not working properly. There were errors in reading files. Though mySQL was connected, MatLab couldn't produce any results as it was not connected with java librabies. We used the source code from github [19] to make Puka work. After importing the JMatLink code from GitHub, the Puka was able to read the file but wasnot able to extract any values from ECG and Respiration Values. Libraries in MatLab and util were imported in the java file to make it work. For example, we added ECG peak detecting libraby in MatLab so that it can count the number of R-R peaks from the particular time frame which we used as input.

The input file was in text format with two columns. The first column is respiration data and te second column is ECG data. The preference file specifies which column holds which data. The JMatLink adopter uses eclipse IDE for java. It has a preference java file and text file which is edited according to our dataset. Some import files are added to make it work properly. The preference file has a clipping id. Those id is used when sniping input data. The minimum length of the input file is 20 seconds and the maximum length of the input data is 10 minutes. The time frame is specified in the java code. To get this java code working we need three major files. A header file, a .ecg file and a .dat file. The header file includes the header details of the input source file, for example the voltage and the time frame for recording. The .ecg file is the ecg signal of each subject. The .dat file contains important information about the software. The .dat files has the configuration information of the fantasia data which can be executed through Puka. Mostly .dat files are text file but the .dat files of fantasia data couldn't be read even after converting it to asimple text file.

To work Puka requires these files. The time content of these files should be equal to the input file provided in Puka. In the java file named "respiration.java", we provide the text file name (input file) and the time frame in several places. We have to mention the time segment we are using. This time segment should match with any of the clipping segment. Then "frmMain.java" is executed in eclipse IDE. We used text file as our input. While loading we have to select the clipping segment depending upon the length of our input file.

Those data for 40 subjects are analyzed in 3 different ways to understand and predict the differences between male – female and old – young.

The variables used in this analyzation are ‘TiMean’, ‘TeMean’, ‘TiStd’, ‘TeStd’, ‘BR’, ‘ShBr’, ‘LgBr’, ‘RRMean’, ‘RRStd’, ‘HR’, ‘ShHr’, and ‘LgHr’. These are calculated from Puka. The variables are explained below.

Table 1: Variable Table

Variable Name	Variable Meaning
TiMean	Mean of inspiration time for a subject in a particular time frame
TeMean	Mean of expiration time for a subject in a particular time frame
TiStd	Standard deviation of inspiration time for a subject in a particular time frame
TeStd	Standard deviation of expiration time for a subject in a particular time frame
BR	Breathing Rate of a Subject at a particular time frame
ShBR	Shortest breath of a subject at a particular time frame
LgBr	Longest breath of a subject at a particular time frame
RRMean	Mean of R-R Interval of a subject in a particular time frame
RRStd	Standard deviation of R-R Interval of a subject in a particular time frame
HR	Heart Rate of a Subject at a particular time frame
ShHr	Shortest beat of a subject at a particular time frame
LgHr	Longest beat of a subject at a particular time frame

The input for puka is respiration and ecg measurement of each subject. The steps to get the RSV and HRV of each subject is shown below.

- Calculate RSV for each subject

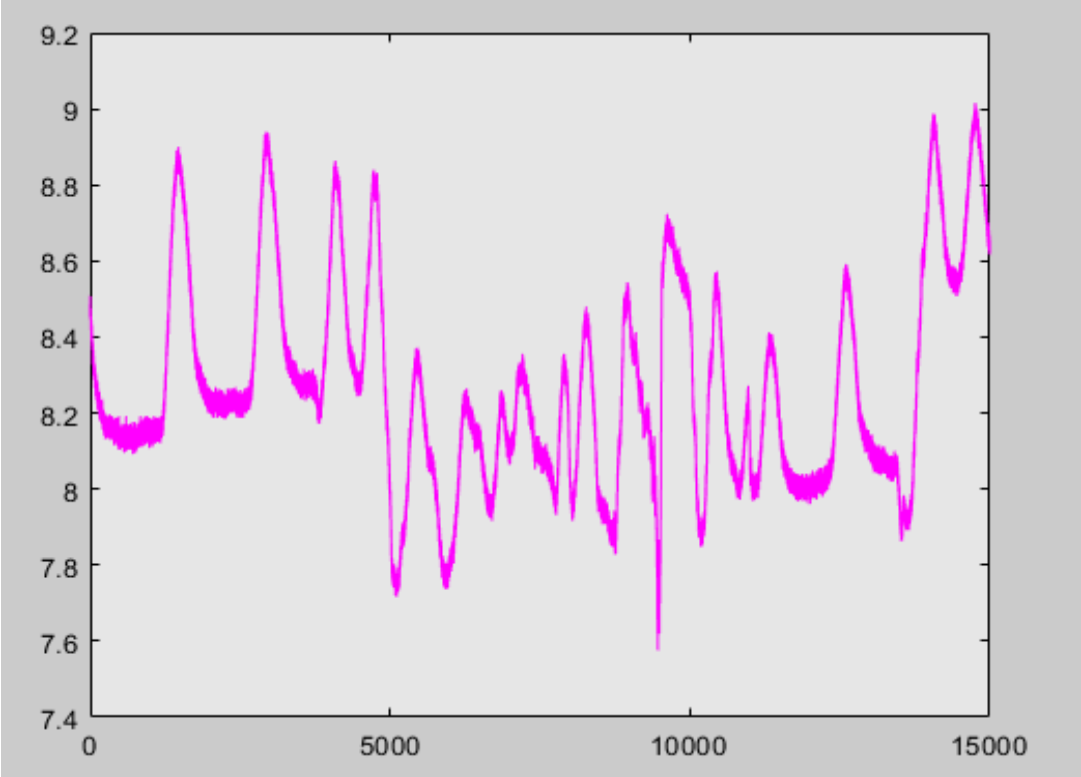

Figure 6: Respiration Graph

a) respiration graph for one young male (1 min data)

b) X-axis – Time (nm)

c) Y-axis – voltage amplitude(mv)

- Detect Peak – Valley from respiration graph

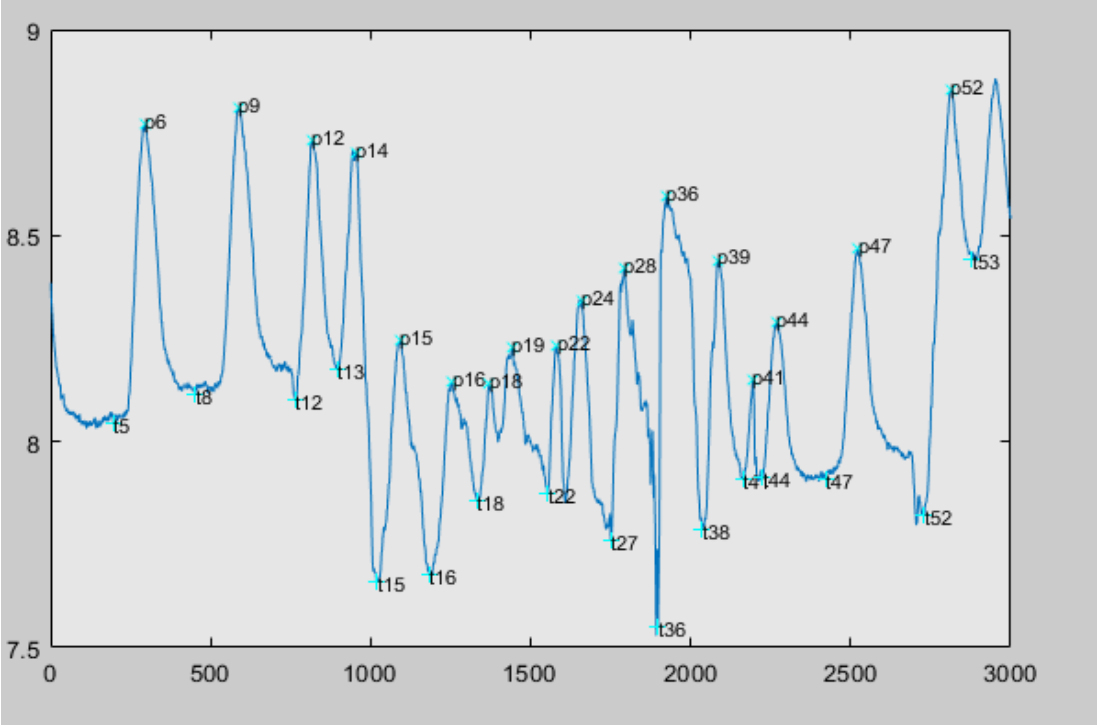

Figure 7: Peak – Valley Detected Graph

a) Corrected peak and valley.

b) Only valid peak and valley considered

c) X-axis – Time (nm)

d) Y-axis – voltage amplitude(mv)

- Calculation of Respiratory Rate Variability from Puka

Table 2: RSV for one subject

Ttotal - mean	seconds	3.365
Ttotal - std dev	seconds	1.6043
Ti - mean	seconds	2.0812
Ti - std dev	seconds	3.6549
Te - mean	seconds	5.4463
Te - std dev	seconds	3.2851
Insp Duty Time		0.6185
RR - mean	cycles/minute	71.3224
RR - std dev	cycles/minute	1.6043
num breaths		17
shortest breath	seconds	0.2
longest breath	seconds	6.24

- Calculate HRV for each subject

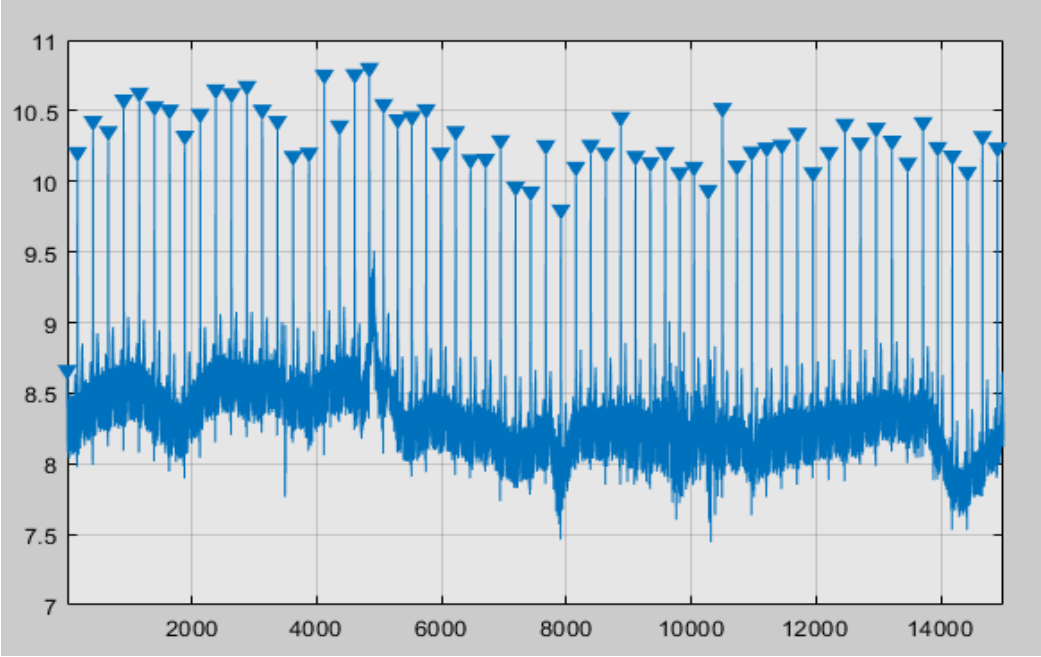

Figure 8: ECG Plot

a) ECG Graph

b) R-R peaks detected

c) X-axis – Time (nm)

d) Y-axis – voltage amplitude(mv)

- Calculate HRV

Table 3: HRV for one subject

total number of R peaks	310
shortest beat (sec)	0.668
longest beat (sec)	1.064
RR interval mean (sec)	0.9671
RR interval std dev (sec)	0.0425
mean instantaneous heart rate (beats/min)	0.0425
IHR std dev (beats/min)	0.0425

5.3 METHODOLOGY

The first step is to determine which features to analyze. This is an important step as the dataset is very small. The data set has 40 subjects. Among them 50% are old, 50% are young and 50% are female and 50% are male. The data set is first divided in 2 groups.

To normalize the data:

1. Determine the *range* = max – min for all value of a feature (eg HR for all data points)
2. Calculate the mean for each sub-group (eg. Mean HR for female and mean HR for male)
3. Calculate the absolute mean difference $diff_{mean}$. (eg. | Mean HR for female – mean HR for male|)
4. The normalized difference is:

$$\frac{100 * diff_{mean}}{range}$$

We considered the feature with the highest normalized difference. We did not consider the intermediate percentages as these may not give reliable results due to the small data size.

5. The normalized values of a feature or variable are determined as follows:

$$\frac{100(Variable\ value - \min value)}{range}$$

5.3.1 Single and Multi-Factor ANOVA

An ANOVA test is a way to find out if survey or experiment results are significant. In other words, ANOV will determine if the null hypothesis is to be rejected or not. In our analysis with single factor ANOVA we divided the dataset into two groups. The age group consisted of old and young people. Young fell in the age range 20-40 and old people fell in range 60-80. The other is the

gender group (male and female). All the subjects in the fantasia database are in good health with no record of any heart or health condition.

5.3.2 Large Group Single Factor ANOVA Analysis

A: Male vs Female – HR Analysis

We looked at the HR feature because based on the literature we expect to see a difference between male and female.

Null hypothesis: HR is same for male and female in any age group with the confidence level 95%.

The analysis is done on male-female using the HR variable. The null hypothesis is rejected with a confidence level of 99%. This says that there is a difference between male HR and female HR.

B: Male vs Female – BR Analysis

We looked at the BR feature because based on the literature we expect to see a difference between male and female.

Null hypothesis: BR is same for male and female in any age group with the confidence level 95%.

The analysis is done on male-female on the BR variable. The null hypothesis is rejected with a confidence level of 99%. This says that there is a difference between male BR and female BR.

C: Old vs Young – HR Analysis

We looked at the HR feature based on the literature we expect to see a difference between young and old.

Null hypothesis: HR is same for old and young in any gender group with the confidence value 95%.

The analysis is done on old-young on the HR variable. The null hypothesis not rejected. This claims that there is a no difference between HR of old and young. This result suggests that in healthy people there is no significant difference between old-young HR. This is a very small dataset and further analysis using a larger dataset is needed.

D: Old vs Young – BR Analysis

We looked at the BR feature based on the literature we expect to see a difference between male and female.

Null hypothesis: BR is same for old and young in any gender group with the confidence value 95%.

The analysis is done on old-young on the BR variable. The null hypothesis does not reject with a confidence level of 99%. This says that there is no difference between male BR and female BR.

The analysis is done on the male female on the BR variable. The null hypothesis is not rejected. This claims that there is a no difference between BR of old and young. This result suggests that in

healthy people there is no significant difference between male-female BR. This is a very small dataset and further analysis using a larger dataset is needed.

E: Male vs Female – RRMean Analysis

We looked at the RRMean feature because based on the literature we expect to see a difference between male and female.

Null hypothesis: RRMean is the same for male and female in any age group with the confidence level 95%.

The analysis is done on male-female on the RRMean variable. The null hypothesis is rejected with a confidence level of 99%. This says that there is a difference between male RRMean and female RRMean.

F: Old vs Young – RRMean Analysis

We looked at the RRMean feature because based on the literature we expect to see a difference between old and young.

Null hypothesis: RRMean is the same for old-young in any age group with the confidence level 95%.

The analysis is done on old-young on the RRMean variable. The null hypothesis is not rejected with a confidence level of 99%. This says that there is no difference between old RRMean and young RRMean.

HR is the inverse of RRMean (RRMean is the mean interval). This is in agreement with the results for HR. This confirms the validity of the dataset.

5.3.3 Sub-Group Single Factor ANOVA Analysis

We further divided the data into 4 sub-groups - (a) old female and young female, (b) old male and young male, (c) old female and old male and (d) young female and old male.

A: Old Female vs Young Female – HR Analysis

We considered the **HR** feature as it had the **highest mean difference**.

Null hypothesis: there is no difference in HR of old female and young female with confidence level 95%.

This hypothesis is not rejected. Therefore, there is no difference in HR between old female and young female.

B: Old Female vs Young Female - BR Analysis

We also considered the **BR** feature as it had the **lowest mean difference**.

Null hypothesis: there is no difference in BR of old female and young female with confidence level 95%.

The hypothesis is not rejected. There is therefore no difference in BR between old female and young female.

C: Old Male vs Young Male - SB Analysis

In the sub-group old male vs young male**, Short Breath and HR** has the **highest and lowest** difference respectively. An analysis is done on both for this group.

Null hypothesis: there is no difference in short breath of old male and young male with the confidence value 95%.

The hypothesis is not rejected. Therefore, there is no difference in Shortest Breath between old male and young male.

D: Old Male vs Young Male - HR Analysis

Null hypothesis: there is no difference in HR of old male and young male with confidence level 95%.

The hypothesis is not rejected. Therefore, there is no difference in HR between old male and young male.

E: Old Female vs Old Male – R-RMean Analysis

In the sub-group old female vs old male, **R-RMean** has the highest difference and **TeStd** has the **lowest difference**. An analysis is done on both for this group.

Null hypothesis: there is no difference in R-RMean of old female and old male with confidence level 95%.

The hypothesis is not rejected reject. Therefore, there is no difference in R-RMean between old female and old male.

F: Old Female vs Old Male – TeStd Analysis

Null hypothesis: there is no difference in TeStd of old female and old male with confidence level 95%.

The hypothesis is not rejected. Therefore, there is no difference in TeStd between old female and old male.

G: Young Female vs Young Male – HR Analysis

In the sub-group young female vs young male**, HR** has **the highest mean difference** and **Long Breath** has the **lowest difference**. An analysis is done on both for this group.

Null hypothesis: there is no difference in HR of young female and young male with confidence level 95%.

The hypothesis is rejected. Therefore, there is a difference in HR between young female and young male.

H: Young Female vs Young Male – LB Analysis

Null hypothesis: there is no difference in long breath of young female and young male with confidence level 95%.

The hypothesis is not rejected. Therefore, there is no difference in long breath between young female and young male.

5.3.3.1 Results

In the sub-group observation, it is noticeable that in each group if the variable of highest difference is in HRV group then the variable with lower difference is in RSV group and vice versa. For the sub-group (a), (c) and (d) a HRV variable has highest mean difference. On the other hand, for sub-group (b) has RSV variable as a highest mean difference. This is calculated using the normalized formula.

The above results show that there is a significant difference between young female HR and young male HR. On the other hand, for this group says that there is no significant difference in BR. The reasons for this is not clear.

We took 2 groups, male vs female and old vs young. The expected result was to get some significant different in these group for their HRV and RSV. The data is normalized as described earlier and a simple analysis is done.

5.3.4 Large Group Single Factor ANOVA Analysis based on mean difference

5.3.4.1 Group 1: Male vs Female

In this group the highest mean difference in the RSV variable group is BR and in the HRV variable group is HR. The lowest difference in this group for RSV variable is breath difference and for HRV variable, shortest beat.

A. RSV variable – BR analysis

Null hypothesis: there is no difference in BR of male and female with confidence level 95%.

The hypothesis is rejected. Therefore, there is a difference in BR between male and female.

B. RSV variable – BD analysis

Breath difference (BD) is the variable which has lowest mean difference in the respiratory data.

Null hypothesis: no difference in breath difference (BD) in male and female with confidence level 95%.

The hypothesis is not rejected and therefore there is no difference in Breath Difference (BD) between male and female.

C. HRV variable – HR analysis

The HR variable has the highest mean difference in the heart data.

Null hypothesis: there is no difference in HR of male and female with confidence level 95%.

The hypothesis is rejected and there is therefore a difference in HR between male and female.

D. HRV variable – SB analysis

The shortest beat variable has the lowest mean difference in the respiratory data.

Null hypothesis: there is no difference in shortest beat of male and female with confidence level 95%.

The hypothesis is not rejected and therefore there is no difference in Shortest Beat between male and female.

5.3.4.2 Group 2: Old vs Young

In this group the highest difference in the RSV variable group is TiStd and in the HRV variable group it is R-RStd. The lowest difference in this group for RSV variable is shortest breath and for HRV variable, longest beat.

A. RSV variable – TiStd analysis

Null hypothesis: there is no difference in TiStd of old and young with confidence level 95%.

The hypothesis is not rejected and therefore there is no difference in TiStd between old and young.

B. RSV variable – SB analysis

Null hypothesis: there is no difference in shortest breath of old and young with confidence level 95%.

The hypothesis is not rejected and therefore there is no difference in Short Breath between old and young.

C. HRV variable – R-RStd analysis

Null hypothesis: there is no difference in R-RStd of old and young with confidence level 99%.

The hypothesis is rejected and therefore there is a difference in R-RStd between old and young.

D. HRV variable – LB analysis

Null hypothesis: there is no difference in longest beat of old and young with confidence level 95%.

The hypothesis is not rejected and therefore there is no difference in longest beat between old and young.

In the old vs young analysis most of cases failed to reject the hypothesis except for R-RStd variable. On the other hand, in male vs female group the highest difference variables rejected the hypothesis. That states that there are differences in healthy male and healthy female HR and BR. As HR and BR rejected the hypothesis in the mal- female group individually, we combined those two variables to do a multifactor ANOVA. As mentioned earlier, since the scale of HR and BR is different, the scales were normalized using

$$\frac{100(Variable\ value - \min value)}{range}$$

5.3.5 Multi-Factor ANOVA

A. HR - BR analysis

The null hypothesis in the male vs female is rejected for HR and BR. We combined these variables for a multi-factor ANOVA.

Null hypothesis: When considering both HR and BR, there is no difference males and females with confidence level 95%.

The null hypothesis is rejected. Hence there is a difference in the HR and BR values between male and female.

B. R-RStd - TiStd analysis

The null hypothesis in the old vs young is rejected for R-RStd but not rejected for TiStd. We combined these variables and did a multi-factor ANOVA.

Null hypothesis: When considering Tistd and RRStd, there is no difference butween old and young with confidence level 95%.

The Null hypothesis is not rejected: This means TiStd and R-RStd considered together do not show any significant difference between young and old.

5.4 SUMMARY OF RESULTS

Table 4: Summary of Results

	Type of ANOVA	Variable	Notes	Null Hypothesis Rejected	F	F-crit	P-value
1	Single Factor	*HR* of Male and Female	Based on literature review	Y	14.41	2.84	0.0005
2	Single Factor	*BR* of Male and Female	Based on literature review	Y	5.16	2.84	0.028
3	Single Factor	*HR* of Old and Young	Based on literature review	N	1.46	7.35	0.23
4	Single Factor	*BR* of Old and Young	Based on literature review	N	0.08	7.35	0.76
5	Single Factor	*RRMean* of Male and Female	Based on literature review	Y	13.41	4.098	0.0008
6	Single Factor	*RRMean* of Old and Young	Based on literature review	N	1.85	4.098	0.181
7	Single Factor	*HR* of Old Female and Young Female	HR has highest mean difference	N	2.86	8.28	0.107
8	Single Factor	*BR* of Old Female and Young Female	BR has lowest mean difference	N	0.0014	4.41	0.96
9	Single Factor	*Shortest Breath* of old male and young male	Shortest Breath has highest mean difference	N	0.029	8.28	0.86
10	Single Factor	*HR* of old male and young male	HR has lowest mean difference	N	0.068	4.41	0.79

11	Single Factor	*R-RMean* of old female and old male	R-RMean has the highest mean difference	N	3.35	8.28	0.08
12	Single Factor	*TeStd* of old female and old male	TeStd has lowest mean difference	N	0.046	4.41	0.83
13	Single Factor	*HR* of Young Female and Young Male	HR has highest mean difference	Y	17.83	8.28	0.0005
14	Single Factor	*Long Breath* of Young Female and Young Male	Long Breath has the lowest mean difference	N	0.038	4.413	0.84
15	Single Factor	*BR* of Male and Female	BR has highest mean difference	Y	5.16	4.1	0.029
16	Single Factor	*BD* of Male and Female	*BD* - has lowest mean difference variable	N	0.043	4.1	0.83
17	Single Factor	*HR* of Male and Female	HR has highest mean difference	Y	14.41	4.1	0.0005
18	Single Factor	*Shortest Beat* of Male and Female	Shortest Beat has lowest mean difference	N	0.14	4.09	0.707
19	Single Factor	TiStd of Young and Old	TiStd highest has mean difference variable	N	2.29	4.098	0.138
20	Single Factor	*Short Breath* of Young and Old	Short Breath has lowest mean difference variable	N	1.5	4.09	0.99
21	Single Factor	RRStd of Young and Old	Choose *RRStd* highest mean difference variable	Y	4.766	4.098	0.035
22	Single Factor	Longest Beat of Young and Old	Choose Longest Beat lowest mean difference variable	N	0.618	4.098	0.436
23	Multi Factor	*BR* of Male and Female *HR* of Male and Female	Null hypothesis rejected for HR and BR (1 and 2) – try both	Y	17.82	3.96	6.65E-05
24	Multi Factor	TiStd of Young and Old RRStd of Young and Old	Null hypothesis rejected for RRStd (6), not for TiStd (5)	N	2.157	3.966	0.145

Figure.7: Analysis

The main results are as follows:

1. HR of Male and Female is significantly different with a confidence level of 95%
2. BR of Male and Female is significantly different with a confidence level of 97.2%
4. HR of Young Female and Young Male is significantly different with a confidence level of 95%
5. RRStd of Young and Old is significantly different with a confidence level of 96.5%
6. Considering both HR and BR of Males and Females, there is a significant difference with a very high confidence level (almost 100%)

The gender can therefore be accurately predicted using HR and BR. The only variable that shows a difference in relation to age is RRStd. Heart information does not seem to be helpful in this regard.

CHAPTER VI

CONCLUSIONS

The main result from this study is that gender can be accurately predicted using HR and BR. The only variable that shows a difference in relation to age is RRStd. Heart information does not seem to be helpful in this regard. Our analysis shows that few of the variables provide significant differences for HRV and RSV between different age group and different gender.

The dataset we used is very small and this may have limited the detection of the contribution of other variables. A bigger dataset is required for future studies. All the subjects are healthy and this may have also caused the minimal differences.

Other machine learning techniques such K-Means clustering can be applied to form multiple age groups. Support Vector Machines, Random Forests, Naïve Bayes and Logistic regression could also be applied, but would require larger datasets.

A control group would have allowed the separation of subjects into health or unhealthy groups for example.

REFERENCES

1. Ken Umetani, Donald H Singer, Rollin McCraty and Mike Atkinson, “Twenty-Four Hour Time Domain Heart Rate Variability and Heart Rate: Relations to Age and Gender Over Nine Decades”, Journal of the American College of Cardiology, Volume 31, Issue 3, March 1998, DOI: 10.1016/S0735-1097(97)00554-8 , http://www.onlinejacc.org/content/31/3/593.abstract

2. Massimo Stefano Silvetti, FabrizioDrago, PietroRagonese, “Heart rate variability in healthy children and adolescents is partially related to age and gender”, International Journal of Cardiology, Volume 81, Issues 2–3, December 2001, doi.org/10.1016/S0167-5273(01)00537-X

3. Jeanne Y. Wei, “Age and the Cardiovascular System”, The New England Journal of Medicine, December 10, 1992, DOI: 10.1056/NEJM199212103272408

4. I A O'Brien, P O'Hare, R J Corrall, “Heart rate variability in healthy subjects: effect of age and the derivation of normal ranges for tests of autonomic function”, Volume 55, Issue 4, dx.doi.org/10.1136/hrt.55.4.348

5. Wilmer W.Nichols, Michael F.O'Rourke,Albert P.Avolio, ToshioYaginuma, Joseph P.Murgo, Carl J.Pepine, C.RichardConti, “Effects of age on ventricular-vascular coupling”, The American Journal of Cardiology, Volume 55, Issue 9, 15 April 1985, doi.org/10.1016/0002-9149(85)90659-9

6. IvanaAntelmi, Rogério SilvaDe Paula, Alexandre R.Shinzato, Clóvis AraújoPeres, Alfredo JoséMansur, Cesar JoséGrupi, " Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease ", The American Journal of Cardiology, Volume 93, Issue 3, 1 February 2004, doi.org/10.1016/j.amjcard.2003.09.065

7. Hulegar A. Abhishekh, Palgun Nisarga, Ravikiran Kisan, Adoor Meghana, Sajish Chandran, Trichur Raju, Talakad N. Sathyaprabha, " Influence of age and gender on autonomic regulation of heart ", Journal of Clinical Monitoring and Computing, Volume 27, Issue 3, June 2013, 10.1152/ajpheart.00903.2005

8. "Effects of aging on the respiratory system", PubMed, May 1984

9. MBehanA. GZabkaG.SMitchell, "Age and gender effects on serotonin-dependent plasticity in respiratory motor control", Respiratory Physiology & Neurobiology, Volume 131, Issues 1–2, July 2002, oi.org/10.1016/S1569-9048(02)00038-1

10. Craig A.Harms, " Does gender affect pulmonary function and exercise capacity? ", Respiratory Physiology & Neurobiology, Volume 151, Issues 2–3, 28 April 2006, doi.org/10.1016/j.resp.2005.10.010

11. Michelle A.CareyJeffrey W.CardJames W.VoltzSamuel J.ArbesJrDori R.GermolecKenneth S.KorachDarryl C.Zeldin, "It's all about sex: gender, lung development and lung disease", Volume 18, Issue 8, October 2007, doi.org/10.1016/j.tem.2007.08.003

12. "Puka - Software for Detection of Breaths in Strain Gauge Recordings", Dec. 15, 2004. Version: 1.0.0 , https://physionet.org/content/Puka/1.0.0/ , Etzel JA, Johnsen EL, Dickerson

JA, Adolphs R. A program to accurately identify peaks in respiration and EKG signals for use in psychophysiological research. Psychophysiology 41 (s1), S73 (2004).

13. “Fantasia Database (fantasia)”, https://archive.physionet.org/cgi-bin/atm/ATM

14. Fantasia Database, 2000 (June 13), https://archive.physionet.org/physiobank/database/fantasia/

15. JP Janssens, JC Pache, LP Nicod, “Physiological changes in respiratory function associated with ageing”, European Respiratory Journal, 1999, PubMed, vol. 13

16. Elaine T. Stathopoulos and Christine Sapienza, “Respiratory and Laryngeal Function of Women and Men During Vocal Intensity Variation”, Journal of Speech Language and Hearing Research, Research Article : 1 Feb 1993, doi.org/10.1044/jshr.3601.64

17. JMatLink Plugin: http://jmatlink.sourceforge.net/docs/ESS99.pdf - JMATLINK

18. GitHub Link for JMatLink Adopter https://github.com/petrickow/INF5960_JMatLinkAdapter

APPENDIX I

ANOVA RESULTS

ANOVA Analysis result: The numbering is done according to the chart in the conclusion

Analysis 1:

Anova: Single Factor	**Heart Rate**					
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1 (Male)	20	1144.8	57.24	63.552		
Column 2 (Female)	20	1351	67.55	83.9721053		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1062.961	1	1062.961	14.4106754	**0.00051446**	2.84244244
Within Groups	2802.958	38	73.7620526			
Total	3865.919	39				

Analysis 2:

Anova: Single Factor	**Breathing Rate**					
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1 (Male)	20	233.2	11.66	7.61726316		
Column 2 (Female)	20	269.4	13.47	5.08010526		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	32.761	1	32.761	5.16028187	**0.02885667**	2.84244244
Within Groups	241.25	38	6.34868421			
Total	274.011	39				

Analysis 3:

Anova: Single Factor	**Heart Rete**					
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1 (Old)	20	1210	60.5	104.606316		
Column 2 (Young)	20	1285.8	64.29	91.3030526		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	143.641	1	143.641	1.46640256	**0.23339185**	7.35254463
Within Groups	3722.278	38	97.9546842			
Total	3865.919	39				

Analysis 4:

Anova: Single Factor	**Breathing Rate**					
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Column 1 (Old)	20	253.8	12.69	6.68621053		
Column 2 (Young)	20	248.8	12.44	7.70252632		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.625	1	0.625	0.0868735	0.76979441	7.35254463
Within Groups	273.386	38	7.19436842			
Total	274.011	39				

Analysis 5:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
HR(Old Female)	10	642.4	64.24	92.3071111		
HR(Young Female)	10	708.6	70.86	60.6204444		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	219.122	1	219.122	2.86569676	0.10772254	8.28541956
Within Groups	1376.348	18	76.4637778			
Total	1595.47	19				

Analysis 6:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
BR(Old Female)	10	134.5	13.45	6.89166667		
BR(Young Female)	10	134.9	13.49	3.83211111		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.008	1	0.008	0.00149201	0.96961331	4.41387342
Within Groups	96.514	18	5.36188889			
Total	96.522	19				

Analysis 7:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
ShortBreath(Old Male)	10	22.12	2.212	0.45797333		
ShortBreath(Young Male)	10	22.76	2.276	0.93322667		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.02048	1	0.02048	0.02944221	0.86567627	8.28541956
Within Groups	12.5208	18	0.6956			
Total	12.54128	19				

Analysis 8:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
HR(Old Male)	10	567.6	56.76	97.4448889		
HR(Young Male)	10	577.2	57.72	36.2084444		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.608	1	4.608	0.06895451	0.79584703	4.41387342
Within Groups	1202.88	18	66.8266667			
Total	1207.488	19				

Analysis 9:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
RRMean(Old Female)	10	9.5167	0.95167	0.0187997		
RRMean(Old Male)	10	10.837	1.0837	0.03315462		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0871596	1	0.0871596	3.3552405	0.0835927	8.28541956
Within Groups	0.4675888	18	0.02597716			
Total	0.55474841	19				

Analysis 10:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
TeStd(Old Female)	10	13.3598	1.33598	0.90669829		
TeStd(Old Male)	10	12.3024	1.23024	1.51069443		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.05590474	1	0.05590474	0.04625209	0.83213609	4.41387342
Within Groups	21.7565345	18	1.20869636			
Total	21.8124392	19				

Analysis 11:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
HR(Young Female)	10	708.6	70.86	60.6204444		
HR(Young Male)	10	577.2	57.72	36.2084444		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	863.298	1	863.298	17.8314139	0.00051173	8.28541956
Within Groups	871.46	18	48.4144444			
Total	1734.758	19				

Analysis 12:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
LongBreath	10	113.68	11.368	20.8817067		
LongBreath	10	108.97	10.897	36.1956456		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1.109205	1	1.109205	0.03886673	0.84592196	4.41387342
Within Groups	513.69617	18	28.5386761			
Total	514.805375	19				

Analysis 13:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Male BR	20	827.1845	41.35922	718.0001		
Female BR	20	1178.641	58.93204	478.8486		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	3088.038	1	3088.038	5.160282	0.028857	4.098172
Within Groups	22740.13	38	598.4244			
Total	25828.16	39				

Analysis 14:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
MALE	20	194.65	9.7325	29.51239		
FEMALE	20	187.76	9.388	25.37703		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1.186802	1	1.1868025	0.043243	0.836378165	4.098172
Within Groups	1042.899	38	27.44470776			
Total	1044.086	39				

Analysis 15:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
MALE	20	743.1472	37.15736	409.3896		
FEMALE	20	1266.497	63.32487	540.9319		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	6847.387	1	6847.387	14.41068	0.000514	4.098172
Within Groups	18056.11	38	475.1607			
Total	24903.5	39				

Analysis 16:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
MALE	20	9.8626	0.49313	0.073113		
FEMALE	20	9.324	0.4662	0.02868		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.007252	1	0.007252249	0.142491	0.707916831	4.098172
Within Groups	1.93405	38	0.050896041			
Total	1.941302	39				

Analysis 17:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
TiStd(OLD)	20	36.6701	1.833505	3.04619156		
TiStd(YOUNG)	20	23.2273	1.161365	0.89485245		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.5177218	1	4.5177218	2.29265229	0.13826173	4.09817173
Within Groups	74.8798364	38	1.97052201			
Total	79.3975582	39				

Analysis 18:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
ShortBreath	20	41.54	2.077	0.61870632		
ShortBreath	20	41.56	2.078	0.67212211		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1E-05	1	1E-05	1.5494E-05	0.99687994	4.09817173
Within Groups	24.52574	38	0.64541421			
Total	24.52575	39				

Analysis 19:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
RRStd(OLD)	20	1.0945	0.054725	0.00091661		
RRStd(YOUNG)	20	1.4929	0.074645	0.00074839		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.00396806	1	0.00396806	4.76643616	0.03526591	4.09817173
Within Groups	0.03163505	38	0.0008325			
Total	0.03560311	39				

Analysis 20:

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
LongBeat(OLD)	20	24.304	1.2152	0.04815596		
LongBeat(YOUNG)	20	25.5005	1.275025	0.06762628		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.03579031	1	0.03579031	0.61823486	0.43657999	4.09817173
Within Groups	2.19986244	38	0.05789112			
Total	2.23565274	39				

Analysis 21:

Anova: Two-Factor With Replication						
SUMMARY	MALE- BRN	MALE - HRN	Total			
MALE						
Count	20	20	40			
Sum	827.184466	743.147208	1570.33167			
Average	41.3592233	37.1573604	39.2582919			
Variance	718.000109	409.389575	553.768218			
FEMALE						
Count	20	20	40			
Sum	1178.64078	1266.49746	2445.13824			
Average	58.9320388	63.3248731	61.128456			
Variance	478.848644	540.931905	501.764112			
Total						
Count	40	40				
Sum	2005.82524	2009.64467				
Average	50.1456311	50.2411168				
Variance	662.260635	638.551162				
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample (Condition :	9566.08157	1	9566.08157	17.8208163	6.65487E-05	3.96675978
Columns(Different V	0.18235031	1	0.18235031	0.0003397	0.985343286	3.96675978
Interaction	369.344096	1	369.344096	0.68805741	0.409423747	3.96675978
Within	40796.2344	76	536.792558			
Total	50731.8424	79				

Analysis 22

Anova: Two-Factor With Replication						
SUMMARY	TiStd	RRStd	Total			
OLD						
Count	20	20	40			
Sum	36.6701	1.0945	37.7646			
Average	1.833505	0.054725	0.944115			
Variance	3.04619156	0.00091661	2.2957856			
YOUNG						
Count	20	20	40			
Sum	23.2273	1.4929	24.7202			
Average	1.161365	0.074645	0.618005			
Variance	0.89485245	0.00074839	0.73912871			
Total						
Count	40	40				
Sum	59.8974	2.5874				
Average	1.497435	0.064685				
Variance	2.03583482	0.0009129				
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	2.12695464	1	2.12695464	2.15786114	0.14596926	3.96675978
Columns	41.0554513	1	41.0554513	41.6520225	9.2064E-09	3.96675978
Interaction	2.39473522	1	2.39473522	2.4295328	0.12322311	3.96675978
Within	74.9114714	76	0.98567726			
Total	120.488613	79				

APPENDIX II

PUKA INPUT FORMAT, FANTASIA DATASET AND JMatLink

Sample Input Data (FANTASIA):

Subject: f1o01

RSV ECG

8.496 8.040
8.468 8.204
8.480 8.144
8.464 7.988
8.480 8.032
8.484 8.184
8.460 8.168
8.496 8.032
8.456 8.008
8.496 8.156
8.444 8.208
8.508 8.040
8.456 7.992
8.508 8.124
8.444 8.228
8.516 8.068
8.456 7.988
8.476 8.100
8.496 8.220
8.452 8.112
8.504 7.996
8.456 8.100
8.508 8.228
8.456 8.168
8.508 8.004
8.460 8.048
8.508 8.208
8.468 8.168
8.508 8.012

8.476 8.024
8.512 8.188
8.468 8.196
8.512 8.016
8.484 7.996
8.496 8.168
8.484 8.200
8.496 8.044
8.508 8.008
8.488 8.120
8.496 8.204
8.472 8.056
8.504 7.964
8.472 8.092

Preference text file from JMatLink Adopter

Before starting to use puka, set the path for WFDB Directory, WFDB Data Path, EGC path. Here we used eclipse IDE to connect WFDB with Cygwin. Cygwin simplifies the configuration and use of the WFDB software packages. Puka used programs from the WFDB Software Package for ECG peak detection. The ecgpuwave application (documentation at http://www.physionet.org/physiotools/wag/ecgpuw-1.htm) is used to identify the EKG waveforms, from which other programs extract the list of RR interval times and location of each R peak. It is recommended to download the WFDB Software Package as source code and install it using Cygwin. JMatLink, enables Java programs to communicate with MATLAB. The JMatLink Adapter is available in GitHub [18]. Import that adapter through eclipse. Set the path for it in the preference file. Eclipse accepts two columns as input, Respiration and ECG data. Mention which column is referred to which data in the "Preference.txt" file in eclipse before execution. Highlighted areas need to be set properly before execution.

WFDBDirectory|C:/cygwin64/home/staff/wfdb-10.5.0/app/
WFDBDataPath|C:/cygwin64/home/staff/wfdb-10.5.0/data/
ConvertECGPath|C:/Users/staff/eclipse-workspace/INF5960_JMatLinkAdapter-master-test/
SampleFreq|250
SignalUnits|mV
SignalGain|2000
ADCResolution|16
ADCZero|0
Length|0:05:0
ClipName|Happy1: 5 min 00 sec
ClipName|Fear2: 1 min 00 sec
ClipName|Happy3: 10 min 0 sec
ClipName|Fear4: 2 min 10 sec
ClipName|Sad5: 1 min 39 sec
ClipName|Happy6: 2 min 20 sec

ClipName|Sad7: 2 min 20 sec
ClipName|Sad8: 2 min 29 sec
ClipName|Fear9: 1 min 40 sec
ClipName|Happy10: 2 min 42 sec
ClipName|Sad11: 1 min 14 sec
ClipName|Fear12: 2 min 4 sec
ClipName|Fantclip: 0 min 20 sec
ClipLength|138000
ClipLength|101000
ClipLength|122000
ClipLength|130000
ClipLength|99000
ClipLength|140000
ClipLength|140000
ClipLength|149000
ClipLength|100000
ClipLength|162000
ClipLength|74000
ClipLength|124000
ClipLength|1200000
ColEKG|2
ColRespiration|1
dbname|newTest
dbname|subjectData
dbname|EduardoMovies

Change the input file name and the time frame of the input data before executing it through eclipse. Parts are highlighted that need to be changed before execution.

After all the required changes listed above are implemented, execute the “frmMain.java” from eclipse to get PUKA to work.

RSM.Java file : A part of it where time frame was specified

```
        private int RunECGPUWAVE() {
                //sends the record ecgTmp created by ConvertECG to ecgpuwave for waveform analysis
                String strCmd = ""; int intExit = 0;
                //        strCmd = frmPreferences.getInstallPath() + "ecgpuwave -r temp -a atr";

                //strCmd = frmPreferences.getInstallPath()+"ecgpuwave -r f1o01 -a ecg -t 00:03:00";
                strCmd = frmPreferences.getInstallPath()+"ecgpuwave -r f1y01 -a ecg -t 00:10:00";
```

```
                try{
                        String[] cmd = new String[3];
                        cmd[0] = "cmd.exe" ; cmd[1] = " /C start " ;
                        cmd[2] = strCmd;  //put my command into the end so executed as if it was
at a c prompt
                        System.out.println("strCmd in RunECGPUWAVE: " +
cmd[0]+cmd[1]+cmd[2]);
                         Process proc =new ProcessBuilder(strCmd).start();

                        StreamGobbler errorGobbler = new StreamGobbler(proc.getErrorStream(),
"ERROR"); //capture error messages
                        StreamGobbler outputGobbler = new
StreamGobbler(proc.getInputStream(), "OUTPUT");  //capture output messages

    errorGobbler.start(); outputGobbler.start();  //start the readers
    intExit = proc.waitFor();  //get processes' exit value to check for errors
    System.out.println("ExitValue: " + intExit);  //if intExit <> 0 then there was an error
                } catch (Throwable t) { t.printStackTrace(); if( intExit == 0 ){ intExit = 1; } }
                return intExit;
        }

        private int RunRDANN() {
          //read the annotation created by ecgpuwave into an external file
                /* code adapted from:
                http://www.javaworld.com/javaworld/jw-12-2000/jw-1229-traps.html
                When Runtime.exec() won't: Navigate yourself around pitfalls related to the
Runtime.exec() method
                @author Michael Daconta */
                String strCmd = ""; int intExit = 0;
                strCmd = frmPreferences.strWFDBPath + "rdann -r f1y01 -a ecg -t 00:10:00 >" +
frmPreferences.getInstallPath() + "ecgOut.txt";

                //strCmd = frmPreferences.strWFDBPath + "rdann -r " +
frmPreferences.getInstallPath() + "f1o01 >" + frmPreferences.getInstallPath() + "ecgOut.txt";
                try{
                        String[] cmd = new String[3];
                        cmd[0] = "cmd.exe" ; cmd[1] = "/C" ;
                        cmd[2] = strCmd;  //put my command into the end so executed as if it was
at a c prompt

                        Runtime rt = Runtime.getRuntime();
    Process proc = rt.exec(cmd);
```

```
                    StreamGobbler errorGobbler = new StreamGobbler(proc.getErrorStream(),
"ERROR"); //capture error messages
                    StreamGobbler outputGobbler = new
StreamGobbler(proc.getInputStream(), "OUTPUT");  //capture output messages

    errorGobbler.start(); outputGobbler.start();  //start the readers
    intExit = proc.waitFor();  //get processes' exit value to check for errors
    System.out.println("ExitValue: " + intExit);  //if intExit <> 0 then there was an error
            } catch (Throwable t) { t.printStackTrace(); if( intExit == 0 ){ intExit = 1; } }

            return intExit;
        }

        private int RunANN2RR() {
            //use ann2rr from the wfdb code to create an interval series.
            //the -i s8 option has the output in seconds with 8 decimal places - NOT sample
units
            String strCmd = ""; int intExit = 0; int intStopTime = 0;

            //strCmd="ann2rr -r f1o01 -a ecg -i s8 -t 0:0:20 >"+frmPreferences.getInstallPath()
+ "rrOut.txt";
            strCmd ="ann2rr -r f1y01 -a ecg -i s8 -t 00:10:00
>"+frmPreferences.getInstallPath() + "rrOut.txt";
            System.out.println("strCmd in RunANN2RR: " + strCmd);
            try{
                String[] cmd = new String[3];
                cmd[0] = "cmd.exe" ; cmd[1] = "/C" ;
                cmd[2] = strCmd;  //put my command into the end so executed as if it was
at a c prompt
                Runtime rt = Runtime.getRuntime();
    Process proc = rt.exec(cmd);
                StreamGobbler errorGobbler = new StreamGobbler(proc.getErrorStream(),
"ERROR"); //capture error messages
                StreamGobbler outputGobbler = new
StreamGobbler(proc.getInputStream(), "OUTPUT");  //capture output messages
    errorGobbler.start(); outputGobbler.start();  //start the readers
    intExit = proc.waitFor();  //get processes' exit value to check for errors
    System.out.println("ExitValue: " + intExit);  //if intExit <> 0 then there was an error
            } catch (Throwable t) { t.printStackTrace(); if( intExit == 0 ){ intExit = 1; } }

            return intExit;
        }

            private int RunIHR() {
            //use ihr from the wfdb code to create a instantaneous heart rate series
            String strCmd = ""; int intExit = 0; int intStopTime = 0;
```

```
		strCmd = frmPreferences.strWFDBPath + "ihr -r f1y01 -a ecg -t 00:10:00 > " + frmPreferences.getInstallPath() + "f1y01.ihr";
		//strCmd = frmPreferences.strWFDBPath + "ihr -r fantasia/f1o01 -a ecg -f s " + frmLoadData.getStartTime() + " -t s" + frmLoadData.getStopTime() +
		// " > " + frmPreferences.getInstallPath() + "f1o01.ihr";

		// "f1o01 -a ecg -f s" + frmLoadData.getStartTime() + " -t s" + frmLoadData.getStopTime() +
		// " > " + frmPreferences.getInstallPath() + "f1o01.ihr";
		System.out.println("strCmd in IHR: " + strCmd);
		try{
			String[] cmd = new String[3];
			cmd[0] = "cmd.exe" ; cmd[1] = "/C" ;
			cmd[2] = strCmd;  //put my command into the end so executed as if it was at a c prompt
			Runtime rt = Runtime.getRuntime();
    Process proc = rt.exec(cmd);
			StreamGobbler errorGobbler = new StreamGobbler(proc.getErrorStream(), "ERROR"); //capture error messages
			StreamGobbler outputGobbler = new StreamGobbler(proc.getInputStream(), "OUTPUT");  //capture output messages
    errorGobbler.start(); outputGobbler.start();  //start the readers
    intExit = proc.waitFor();  //get processes' exit value to check for errors
    System.out.println("ExitValue: " + intExit);  //if intExit <> 0 then there was an error
		} catch (Throwable t) { t.printStackTrace(); if( intExit == 0 ){ intExit = 1; } }

		return intExit;
	}
```

VITA

IPSITA GHOSH

COMPUTER SCIENCE

Master of Science

Thesis Title: HEART AND RESPIRATION FUNCTION RELATED TO AGE AND GENDER

Major Field: Computer Science

Biographical:

Education:

Completed the requirements for the Master of Science in Computer Science at Oklahoma State University, Stillwater, Oklahoma in July 2019.

Completed the requirements for the Bachelor of Technology in Information Technology, West Bengal University of Technology, West Bengal, India in August 2016.