MULTIVARIATE APPROACH FOR SHOE SIZE MODELLING OF MALAYSIAN WOMEN

MOHD FAIZAL BIN MOHD HAMZAH

UNIVERSITI TEKNOLOGI MALAYSIA

MULTIVARIATE APPROACH FOR SHOE SIZE MODELLING OF MALAYSIAN WOMEN

MOHD FAIZAL BIN MOHD HAMZAH

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia

OCTOBER 2020

DEDICATION

This thesis is dedicated to my wife, who taught me that nothing is impossible if we have the desire to achieve it. It is also dedicated to my parent, who taught me that it is required to accomplish one task at a time in order to complete the largest task.

ACKNOWLEDGEMENT

Assalamualaikum w.b.t and Alhamdullilah, first and above all, I am grateful to Allah, our Lord and Cherisher, the Almighty for providing me this opportunity and granting me the capability to proceed successfully to complete this thesis after went through past few years of hard works, late night work or early mornings into the completion of this thesis. I have faced several difficulty, conflict and failures along the journey. This thesis appears in its current form due to the assistance and guidance of several people. I would therefore like to offer my sincere thanks to all of them.

First and foremost, I would like to express the deepest appreciation to my main supervisor Professor Dr Norliza Mohd Noor, who has shown the attitude and the substance of a genius: she continually guide in regard to do actual research, and an excitement in regard to teaching. Without her supervision and constant help this thesis would not have been possible.

I would like to thank to Dr Omar Mohd Rijal the insightful discussion, offering valuable advice, for your support during the whole period of this research. In addition, a thank you to Nurkhairany Amyra Mokhtar for your support in collecting the data. I am also indebted to Universiti Teknologi Malaysia (UTM) for providing good facility in supplying the relevant literatures. Many thanks to The Ministry of Education Malaysia (MOE), University of Malaya (UM) and Universiti Teknologi Malaysia (UTM) for funding the research grant for this study.

I cannot finish without thanking my family. I warmly thank and appreciate my parents and my mother and father-in-law for spiritual support in all aspects of my life. I also would like to thank my brothers, sister, and brothers and sisters-in-law, for their assistance in numerous ways. I want to express my gratitude and deepest appreciation to my lovely sons and daughter; Emir Naufal El-Qushayyi, Emir Darwish Elfaruqi and Eisya Faiha El Adzra who’s always give me motivation to move on. And finally, I know that you did not want to be named, my lovely wife, Dr Nor Fahimah Mohd Razif, who have finished her PhD in a young age. Without your supports and encouragements, I could not have finished this work, it was you who kept the fundamental of our family, and I understand it was difficult for you, therefore, I can just say thanks for everything and may Allah give you all the best in return.

Finally, the will to win, the desire to succeed, the urge to reach your full potential, these are the keys that will unlock the door to personal excellence.

ABSTRACT

The problem of finding the best fitting shoes among Malaysian women is due to the absence of a Malaysian women shoe sizing system. Currently, shoe manufacturers in Malaysia use United Kingdom (UK) and United States of America (USA) size which is based on the ISO 9407:2019 standard: Mondopoint shoe sizing system, which is basically developed based on foot length and foot breadth. The existing shoe sizes do not cover the width and girth part that is a unique foot shape for Malaysian women. Improper shoe sizing can interrupt the biomechanics of the foot and ankle causing pain and increasing the risk of falling to the wearer. This may create safety issues especially in the workplace like strained foot muscles, slips and falls, and body fatigue which gradually grow into severe musculoskeletal disorders or rheumatic diseases if they are unattended. These problems and related issues were addressed in this study using statistical modelling of foot measurements for the development of the Malaysian women shoe sizing system. The objectives of this study are to investigate the selection of foot measurements used in modelling foot measurements, to enhance existing clustering technique for modelling shoe size, to model foot size variation with multivariate normal model and piecewise regression, and to develop a process to determine shoe size. A set of foot measurement was measured by Infoot 3-dimensional foot scanner; foot length (FL), foot breadth (FB) and foot's ball girth (BG) were used to represent shoe size. 252 women from University of Malaya (UM) volunteered to have their foot measurements taken using the Infoot 3D foot scanner. The data was then split into 161 training data and 91 testing data. The variations of size were investigated using an improved clustering analysis and the multivariate normal mixture probability distribution where a shoe size model is proposed. The process was repeated to five-foot measurements (FL, FB, BG, instep length (IL), and fibulare instep length (FIL)). As for three-foot measurements, FL is treated independently of FB and BG. Conditional on fixed FL value, subdivisions were identified for selected (FB, BG) combination. Taking advantage of the existence of normal probability distribution, univariate hypothesis testing was performed to test group separation and chi-square test to test goodness fit of subdivisions. An alternative approach called piecewise regression was used for five-foot measurements (FL, FB, BG, IL, FIL) to model shoe size as comparison. The study successfully attained the performance of aggregate loss which is less than benchmark value and coverage percentage of 71.25% for the training data. The proposed shoe size was then verified with testing data and showed high coverage percentage of 82.53%. This study concludes that a modest sample size of 252 women was sufficient to develop a prototype shoe sizing system using the proposed novel approach. In conclusion, this study provides a novel process to create a Malaysian women shoe sizing system that considers ball girth (BG), in which most Malaysian women have unique shape. The proposed system provides a significant improvement to the Mondopoint sizing system for Malaysian women. In addition, it also allows for a test of the significant separation of shoe sizes that is overlooked by most ergonomics by using a multivariate normal mixture model.

ABSTRAK

Masalah dalam mencari kasut yang sesuai di kalangan wanita Malaysia adalah kerana ketiadaan sistem ukuran kasut wanita Malaysia. Pada masa ini, pengeluar kasut di Malaysia menggunakan ukuran United Kingdom (UK) dan Amerika Syarikat (AS) yang mengikut piawai ISO 9407:2019 Sistem ukuran kasut Mondopoint yang pada dasarnya dibangunkan berdasarkan panjang kaki (FL) dan kelebaran kaki (FB). Ukuran kasut sedia ada tidak merangkumi bahagian kelebaran dan lilitan bola kaki yang unik untuk wanita Malaysia. Ukuran kasut yang tidak tepat boleh mengganggu biomekanik kaki dan buku lali sehingga mengakibatkan kesakitan dan meningkatkan risiko jatuh bagi si pemakai. Situasi ini boleh menimbulkan masalah keselamatan terutamanya di tempat kerja seperti ketegangan urat kaki, tergelincir dan jatuh, dan keletihan badan yang boleh mengakibatkan gangguan rangka otot dan penyakit reumatik sekiranya tidak ditangani. Kajian ini menangani masalah dan isu yang berkaitan dengan menggunakan model statistik pengukuran kaki untuk membangunkan sistem ukuran kasut wanita Malaysia. Objektif kajian ini adalah untuk mengkaji pemilihan ukuran kaki yang digunakan dalam model pengukuran kaki, menambah baik teknik kelompok yang sedia ada dalam membentuk model ukuran kasut, membangunkan model berasaskan kepelbagaian ukuran kaki menggunakan model multivariat normal dan regresi *piecewise*, dan seterusnya membangunkan proses penentuan saiz kasut. Satu set ukuran kaki diukur menggunakan mesin pengimbas kaki Infoot 3-dimensi; panjang kaki (FL), kelebaran kaki (FB) dan lilitan bola kaki (BG) digunakan untuk mewakili ukuran kasut. 252 wanita dari Universiti Malaya (UM) secara sukarela melakukan pengukuran kaki mereka menggunakan mesin pengimbas kaki Infoot 3-D. Data tersebut kemudian dibahagikan kepada 161 data latihan dan 91 data ujian. Kepelbagaian ukuran dikaji menggunakan analisis pengelompokan yang ditambah baik dan taburan kebarangkalian campuran multivariat normal di mana model ukuran kasut dicadangkan. Proses ini diulang dengan menggunakan lima ukuran kaki (FL, FB, BG, panjang kekura kaki (IL) dan panjang kekura fibular (FIL)). Bagi tiga ukuran kaki, FL diukur secara berasingan dari FB dan BG. Bagi nilai FL yang tetap, sub-bahagian ditentukan bagi kombinasi (FB, BG) terpilih. Berdasarkan kewujudan taburan kebarangkalian normal, ujian hipotesis univariat dijalankan untuk menguji pemisahan kumpulan dan ujian kuasa-dua untuk menguji kesesuaian sub-bahagian. Pendekatan alternatif yang disebut regresi *piecewise* digunakan untuk lima ukuran kaki (FL, FB, BG, IL, FIL) untuk model ukuran kasut sebagai perbandingan. Kajian ini berjaya mencapai prestasi kerugian agregat yang kurang daripada nilai penanda aras dan peratusan liputan sebanyak 71.25% untuk data latihan. Ukuran kasut yang dicadangkan kemudiannya disahkan dengan data ujian dan menunjukkan peratusan liputan yang tinggi iaitu 82.53%. Kajian ini menyimpulkan bahawa ukuran sampel sederhana dari 252 wanita memadai untuk membentuk prototaip sistem ukuran kasut menggunakan pendekatan baru yang dicadangkan. Kesimpulannya, kajian ini menyediakan proses baru dalam membentuk sistem ukuran kasut wanita Malaysia yang mempertimbangkan BG, di mana kebanyakkan wanita Malaysia mempunyai bentuk yang unik. Sistem yang dicadangkan memberikan peningkatan yang signifikan terhadap sistem ukuran Mondopoint untuk wanita Malaysia. Di samping itu, ia juga membolehkan untuk menguji pemisahan saiz kasut yang diabaikan oleh kebanyakan ahli ergonomik dengan menggunakan model multivariat normal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FSS	-	Foot Sizing System
FL	-	Foot Length
FB	-	Foot Breadth
BG	-	Ball Girth
IL	-	Instep Length
FIL	-	Fibulare Instep Length
EM	-	Expectation Maximization
MLE	-	Maximum Likelihood Estimator
LRT	-	Likelihood Ratio Test
AIC	-	Akaike's Information Criterion
BIC	-	Bayesian Information Criterion
CPCC	-	Cophenetic Correlation Coefficient
TEMS	-	Technical Error Measurements
SE	-	Standard Error
MSE	-	Mean Square Error

LIST OF SYMBOLS

d	-	Difference between first and second measurement
N	-	Number of samples
f	-	Probability density distribution
π_r	-	Mixing proportion
$\boldsymbol{\theta_r}$	-	Estimated parameter
$g_r(.)$	-	Component densities of the distribution
r	-	Number of clusters
s	-	Maximum number of clusters
H_0	-	Null hypothesis
H_1	-	Alternative hypothesis
m	-	Number of parameters
μ	-	Mean
σ	-	Standard deviation
Z	-	Expected value
χ^2	-	Chi-square statistic
f_i	-	Observed frequency
e_i	-	Expected frequency
r_{ij}	-	Correlation between foot measurements

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

It is very crucial to choose ideal footwear since it plays a major role in protecting the foot in daily activity. Choosing the right footwear can help ensure the long-term health of our feet, as well as our entire body because it will provide essential mechanical support that upholds body weight and structural balance while withstanding the pressure, most of the time. The best fit for a user may depends on their daily use.

The search for the ideal footwear leads to the need for a foot sizing system (FSS) for Malaysian women. The development of FSS for Malaysian women is from two approaches; firstly statistical methods and ideas using foot measurements to represent the foot shape for examples foot breadth (FB) and foot length (FL), and secondly the ideas of expressing or measuring shape using Fourier Descriptor (FD) (Zhang and Lu, 2004). This thesis focuses on the first approach, and only gives introduction of the second approach. The study on the variation of foot measurement is important since it will determine a standard for shoe size determination.

Malaysia is one of the largest producers of footwear in Asia as stated in the MATRADE's report (MATRADE, 2015). This success is due to the ability to deliver high-quality products at reasonable prices as well as extensive marketing. Besides having exports to South East Asia, a few European and the Middle East countries, most of the footwear is sold to the Malaysian market. The Malaysian footwear industry is continuously growing since it is supported by an experienced and skilled workforce with both technical and practical skills (Insights, 2014). Currently, there are 1,000 footwear manufacturers in Malaysia employing some 30,000 people. They are mainly located in the states of Perak, Selangor, and Johor (Insights, 2014).

These number emphasize the significance of the local shoe-making industry and its contribution to the country's economy. For instance, the industry recorded a stunning RM1 billion in sales in 2013 as reported by the Malaysian Tourism Ministry, involving RM621 million footwear exports to the international market as well as RM300 million in the domestic market (Awani, 2013). As comparison, MATRADE statistics show total footwear exports in 2017 slightly increased to RM652.7 million (Shah, 2018). Malaysia was ranked 19^{th} worldwide regarding shoe exports, with Ipoh as the main footwear production center for the country (Awani, 2013).

A 4% increase in footwear export was recorded in the year 2014 for Malaysian footwear manufacturers, of which 90% was for high-end ladies' footwear (MATRADE, 2015). For Malaysia to move from Asia's leading shoe manufacturer to become a global competitor, Malaysian footwear manufacturers should be proactive to adopt new technology and innovation to improve footwear quality and refine their skills and production. Since ladies' footwear dominates the share of footwear production, it is important to concentrate on this market segment.

There are five type of foot measurement being used this research which is foot length (FL), foot breadth (FB), ball girth (BG), instep length (IL) and fibulare instaep length (FIL). Table 1. 1 shows the detail of the main foot measurement being used in this study. Figure 1. 1 and Figure 1. 2 show illustration of foot measurements of foot length, ball girth, foot breadth, instep length, Fibulare instep length and anatomical bone and landmark names (I-Ware Laboratory Co., 2017).

Table 1. 1 Definition of Foot Measurement Used

Foot Measurement	Definition
Foot Length (FL)	The distance between the back point of the heel and the foremost of the longest toe. Refer Figure 1. 1.
Foot Breadth (FB)	The distance between the metatarsal tibial and metatarsal fibulare of the ball cross section projected to the standing surface. Refer Figure 1. 1.
Ball Girth (BG)	Circumference of foot. It measured around the last from the lateral ball point to the medial ball point. Refer Figure 1. 1.
Instep Length (IL)	The distance between the back point of the heel and the first metatarsophalangeal (MTP) protrusion. Refer Figure 1. 1.
Fibulare Instep Length (FIL)	The distance between the back point of the heel and the fifth metatarsophalangeal (MTP) protrusion. Refer Figure 1. 1.

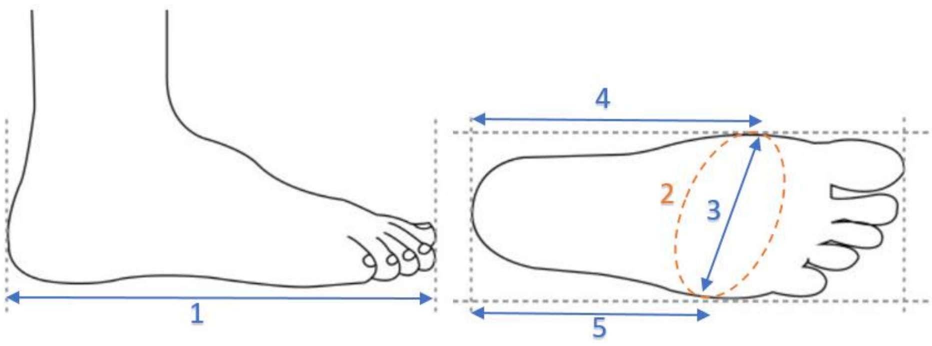

Figure 1. 1 Illustration of Foot Measurements.1:Foot length, 2: Ball girth, 3: Foot breadth, 4: Instep length, 5: Fibulare instep length

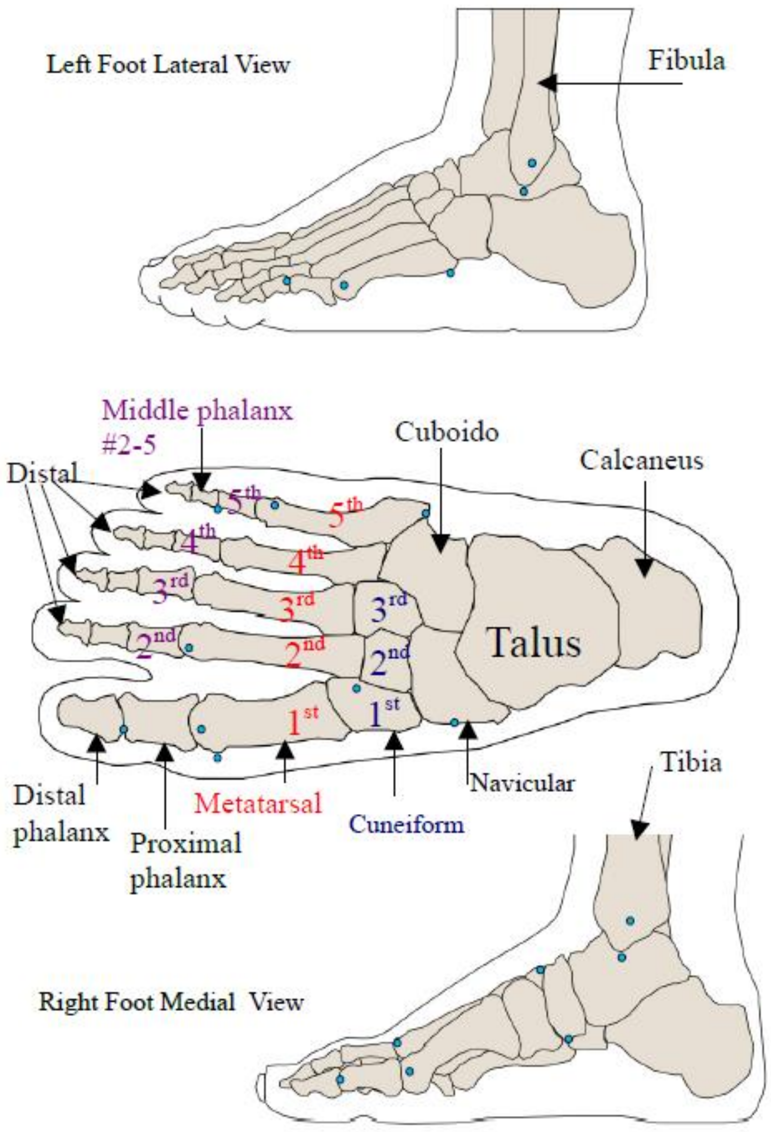

Figure 1. 2 Anatomical Bone and Landmark Names

1.2 Problem Background

The production of shoe was made depends on the shoe last which is a 3-dimensional wooden or plastic mould, that provides human anatomical foot measurement. The shoe last used will determine the overall fit of the shoe and describes the foot shape measurement such as heel width, instep height, forefoot width, and toe box depth. The shoe last was made following the standard sizing system to fit a certain population for the economic and large-scale production purpose in the shoe industry (Hinojo-Pérez *et al.*, 2016). All of this is made possible only if the following problems are considered.

One of the issues in developing a shoe sizing system is a different group of peoples might have a different sizing system for example; the shoe sizing system for adults and children is different. It also differs for men and women (Mishra *et al.*, 2017). The knowledge of sex-related differences in foot measures is vital to assist proper shoe fit in both men and women (Hong *et al.*, 2011; Krauss *et al.*, 2008; Krauss *et al.*, 2010). Nowadays, there are sizing systems developed particularly for activity-specific, function-specific and fashion-specific shoes. At least two categories of shoe sizing standard exist, either based on stick length or actual foot length (Cheng and Perng, 2000). The stick length method categorized the foot based on the measurement of foot length only where it is measured from the last toe to heel top points. The measurement is longer than the shoe last bottom length. It is widely used in European countries, America and Britain such as the British Sizing System, French (Continental) Sizing System, Euro point Sizing System. The second category is based on foot length and foot girth or foot width. Shoe sizing standards that apply this method include the Japanese Sizing System, Mondopoint Sizing System, and Mainland China sizing system.

Another issue concerns health and injury. The improper shoe size selected by the user exerts constant pressure on a certain area of the foot, mostly on the medial and lateral surfaces of the foot that can be produced from shoes that are too small. While the shoes that are too large can cause the foot sliding within the shoes, and the effect is friction ulcers that occur behind the heel (Most and Sinnock, 1983; Reddy *et al.*, 1989). Besides that, improper shoe sizing can interrupt the biomechanics of the foot and ankle causing pain and increasing the chances of falling (Harrison *et al.*, 2007; Manna *et al.*, 2001). This creates safety issues in the workplace like strained foot muscles, slips and falls, body fatigue and other problems (Witana *et al.*, 2004), which gradually grow into severe musculoskeletal disorders or rheumatic diseases if not addressed. A lot of research have been done from different perspectives such as ergonomics, anthropometry, engineering physiology, industrial design, and foot science to address the issue of injury. Some studies enhanced footwear performance and safety and produced a positive impact on the user and market. Figure 1. 3 shows the example of the effect of using the wrong shoe size.

Bunion Blister Ankle sprain

Figure 1. 3 Example of the effect of using the wrong shoe size

A major commercial concern is that of shoe comfort; a problem for both producer and consumer. A shoe's comfort may be influenced by the inside shoe climate and factors such as colour and fashion (Slater, 1986). Other factors that may lead to user satisfaction are foot shape (Hawes *et al.*, 1994), skeletal alignment (Miller *et al.*, 2000) and shoe fit (Witana *et al.*, 2004). Improper fit of footwear may cause slips, falls (Hignett and Masud, 2006) or body fatigue (Lin *et al.*, 2007). Other than that, the material properties influence the shoe's comfort (Goonetilleke, 1999; Mündermann *et al.*, 2002; Yung-Hui and Wei-Hsien, 2005; Zhang *et al.*, 1991).

Several approaches have been introduced to ensure comfortability of shoes produced. Au and Goonetilleke (2007) sought to find the preferred fit in the different regions of the shoe along with the characteristics that distinguish between comfortable and uncomfortable shoes. In the study, the shoes were divided into six regions; the toe region, Metatarso-Phalangeal Joint (MPJ) region, arch region, rearfoot region, Ingress/Egress opening and fastening type. The rear fit region and fastening type do not show a significant difference between different fit ratings. The study concludes that the comfortable shoe does not require a perceived fit in every region of the shoe. The above issues aggravate the shoe design problem where if consider too many foot measurements, it will require a large number of shoe size and is a problem in shoe manufacturing.

Branthwaite *et al.* (2013a) focused on how the toe box shape can affect dorsal and plantar pressure with particular interest around the forefoot in a healthy female population. Three different types of toe box styles were examined; round, squared and pointed. As a result, the shoes with a round toe established the least pressure around the medical aspect of the toe while the pointed shoes distributed least amount of pressure in the lateral toe area. Another finding is that the volume of the foot does not correlate to forefoot pressure. The study also recommends shoe manufacturers or designers to consider footwear design around the toe box to improve fit and reduce pressure.

Even though the consideration of foot function and foot health are important in footwear choice, Branthwaite *et al.* (2013b) consider a different perspective. They claim that footwear choices by younger women are activity-specific and subjects chose the style and design of shoes related to the image they want to portray. The summary for problem background regarding the shoe sizing system shown in Figure 1. 4.

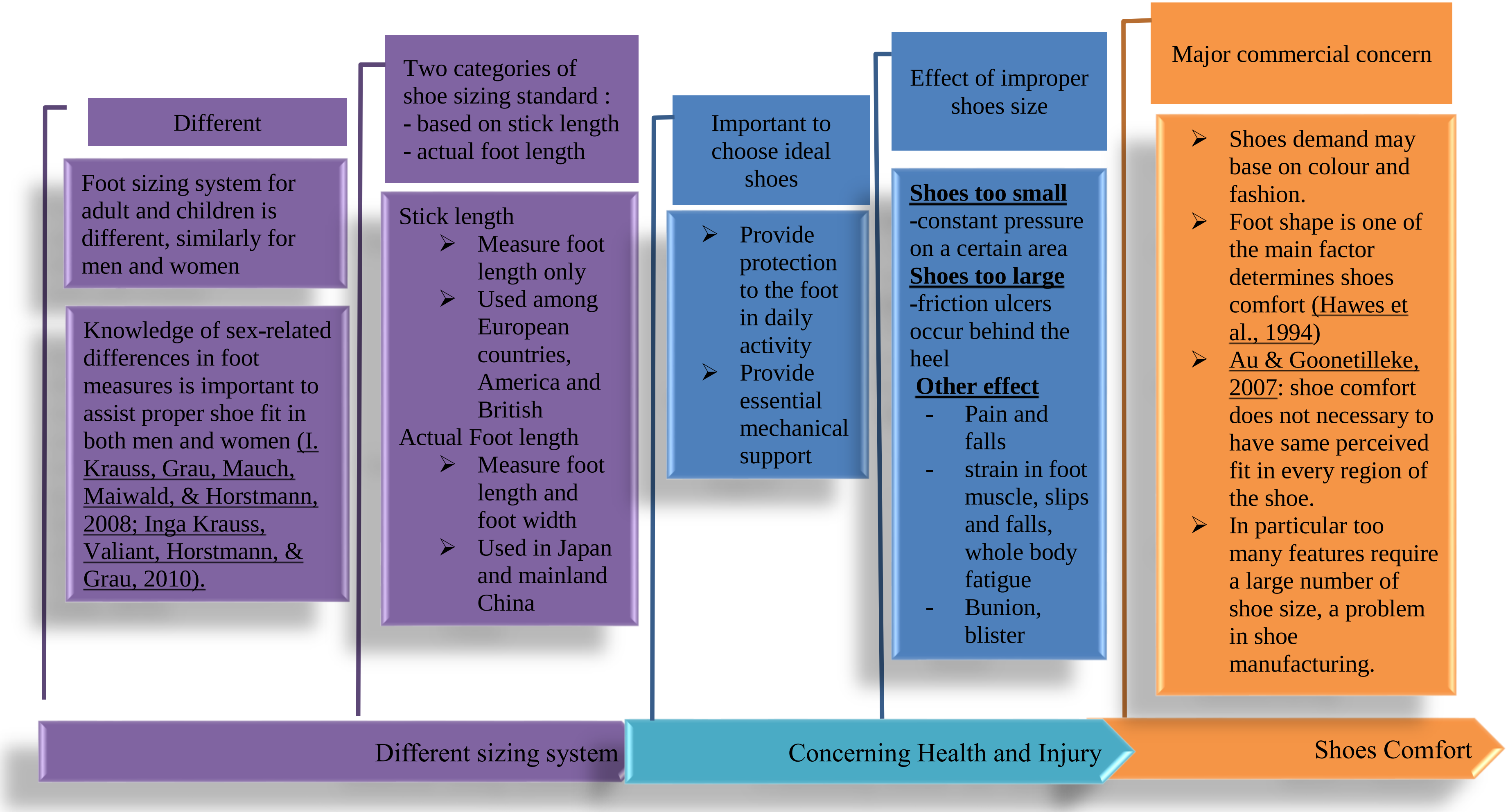

Figure 1. 4 Summary of Problem Background

1.3 Problem Statement

To the best of our knowledge, there is no standard shoe sizing for Malaysian women with little attention paid to model or measure variations in shoe shape. Currently, Malaysia is using the sizing systems of other countries such as ISO 9407:2019 (Standard, 2019) which is the Mondopoint shoe sizing system. Ibrahim and Khedifb (2004) disclosed that most Malaysian women have difficulty choosing the right shoe size as they have a unique foot shape, particularly at the girth and width. This is proven by surveys conducted by Chua *et al.* (2013) and Shariff *et al.* (2014) which revealed that more than 60% of Malaysian women have difficulty choosing the right shoe size and are dissatisfied with the existing selected shoe size. Therefore, it is crucial for Malaysian women to have their shoe sizing system.

This thesis applies statistical methods and ideas in particular multivariate approach to address issues (i) choice of shape features, (ii) the number of shape groups, (iii) classification of subjects using shape features and (iv) general inference from shape features, to determine the foot shape of Malaysian women. Historically, ergonomics and allometry dominate all issues concerning the shape of the human anatomy. Issues (i), (ii), (iii) and (iv) have been addressed using ad-hoc approaches and methods especially on the selection of shape features from statistical views and the separation between different group of shapes (Cheng and Perng, 2000; Hill *et al.*, 2017; Kalebota *et al.*, 2003; Lee and Wang, 2015). This study offers a systematic approach to the issues by using foot measurements to represent foot shape.

Agglomerative hierarchical clustering analysis is commonly used in foot shape classification (Bataller *et al.*, 2001; Lee and Wang, 2015; Mauch *et al.*, 2009). However, different results may be found when using different clustering algorithms or techniques (Everitt, 1981). Because the heuristic method is infrequently conclusive, the model-based clustering technique was developed by Fraley and Raftery (2002) to conform with the standard statistical method and a multivariate normal mixture model. However, it can lead to be over-fitting model if the number of clusters is treated as the number of mixing elements in the multivariate normal mixture model (Baudry *et al.*, 2010). There was no application of the multivariate normal mixture to study the human foot shape variation being done since most of the classification studies satisfied with the outcome from cluster analysis. Thus, this study utilizes the multivariate normal mixture model to enhance the clustering technique for modelling shoe size.

1.4 Research Objectives

This study aims to propose a process that can help create a Malaysian women shoe sizing system based on multivariate approaches. To achieve the aim, this study has four main objectives which are;

(a) To investigate the selection of foot measurements in modelling shoe size.

(b) To enhance the clustering technique for modelling shoe size.

(c) To model foot size variation with multivariate approaches.

(d) To develop a process to determine Malaysian women shoe size.

At the end of the research, a novel process is proposed for the development of a Malaysian women shoe sizing system based on multivariate approaches with a view of improving the existing shoe sizing system.

1.5 Scope of The Research

This study will concentrate on women's shoe sizing for ladies above 18 years only. The final closure of growth plates in the feet normally occurs by the age of 18 years old. Hence, the data was randomly collected for women aged 18 years old and above, from University of Malaya staff and students which represents the population fairly well. The data was collected from the previous study by using a research grant funded by The Ministry of Education Malaysia, University of Malaya and Universiti Teknologi Malaysia (Shariff *et al.*, 2014).

In order to address issues of (i) choice of shape features, (ii) the number of shape groups, (iii) classification of subjects using shape features and (iv) general inference from shape features as stated in Section 1.3, and to achieve the objectives as stated in Section 1.4, this study focused on (a) Modelling foot shape using existing statistical technique, (b) Providing critical appraisal of existing technique using statistical ideas or theory, (c) Investigating choice of foot measurement using clustering of variables, (d) Emphasizing on statistical test for multivariate normal mixture; example are means of 2- ellipsoid different, and (e) Proposing a sizing system by applying subdivision of the ellipsoid technique. The approach undertaken in this study is for small sample size analysis. This study focused on the flat type of shoe sizes.

1.6 Significance of The Research

This study proposed a process that complements the Mondopoint sizing system for creating the appropriate shoe sizing system for Malaysian women. The proposed process allows for statistical test to investigate the choice for (i) number of clusters or groups (ii) the significant subdivision of foot measurements.

1.7 Definition of Key Terms

This study contains certain key terms. Although these key terms are discussed in detail in Chapter 2 and Chapter 3, they are briefly introduced in this early chapter of the thesis to allow the reader to make sense of what is presented in the subsequent Chapters. Table 1.2 shows the definition of few key terms used in this study.

Table 1. 2 Definition of Key Terms

Key Terms	Definition
Technical Error Measurements (TEMS)	Index to express the error margin in anthropometry where it represents the accuracy and the quality of the tool used in this study.
Cophenetic Correlation Coefficient (CPCC)	Correlation between original distance matrix and cophenetic distance matrix. An optimal cluster is defined when the large value of the CPCC index.
Expectation Maximization (EM) algorithm	An iterative method to find maximum likelihood estimates of parameters in statistical models in the presence of latent variables.
Ellipsoid	The geometrically representative of a probability mixture model of 3 foot measurements
Akaike's Information Criterion (AIC)	An estimator of out-of-sample prediction error and thereby relative quality of statistical models for a given set of data. It is an information criteria that going to be used in justify number of clusters.
Bayesian Information Criterion (BIC)	An index used in Bayesian statistics to choose between two or more alternative models. It is also an information criteria that going to be used in justify number of clusters.
Likelihood Ratio Test (LRT)	A test to assesses the goodness of fit of two competing statistical models based on the ratio of their likelihoods, specifically one found by maximization over the entire parameter space and another found after imposing some constraint. It is used to justify the number of clusters gained from EM algorithm.

1.8 Overview of The Thesis

This thesis is divided into five chapters. It starts by introducing the topic in Chapter 1 which includes foot size variations introduction, the background of study and a problem statement. This chapter also explains the objectives, scope of the research and data analysis methods that were used in this research.

Chapter 2 reviews the literature relevant to the study. It begins by reviewing the literature concerning the shoe problem, followed by the problems faced by women in Malaysia concerning foot shape. The last section is an overview of the methods of foot shape modelling including clustering analysis, multivariate normal mixture technique, and piecewise regression technique. Chapter 2 also reviewed the validation techniques applied to the proposed shoe sizing system.

Chapter 3 details the methodology of the research. It consists of flowcharts that help the reader understand the method applied in the research. It also offers an overview of the Agglomerative Hierarchical clustering analysis including distance measure and linkage method used in this research, piecewise regression, and multivariate normal mixture model. This study used four types of distance which were Euclidean, Manhattan (City block), Maximum and Canberra. The linkage methods employed in this study are single, complete, group average, centroid, median and Ward linkage.

Chapter 4 starts with providing the explanatory data analysis including descriptive statistical data analysis, the selection of foot measurement, the accuracy of measurement and outlier detection. It also provides the shoe size modelling study through multivariate normal mixture method including; the selection of distance measure and linkage method based on the Cophenetic Correlation Coefficient (CPCC) index, the optimal number of component based on likelihood ratio test (LRT) p-value, Akaike's information criterion (AIC), and Bayesian information criterion (BIC). Chapter 4 also presents the shoe size modelling study through the piecewise regression method including; the selection y-axis variables and x-axis variables by using the cluster of features analysis; and determining breakpoint for piecewise regression.

Chapter 5 starts with providing a blind shoe sizing system. Due to a large number of shoe sizes, this chapter provides a new shoe sizing system creates from a multivariate normal mixture model. The one-way ANOVA test and post hoc then give how is the separation between shoe sizes created from a multivariate normal mixture model. Then, it also gives a comparison between the shoe sizing system created from 3-foot measurement and 5-foot measurements. Chapter 5 also gives the validation result of the subdivision proposed by a multivariate normal mixture model besides the validation of using the test dataset. Chapter 5 ends with showing the shoe size created from the piecewise regression model as a comparison.

Chapter 6 conclude the novelty of the study, short comings and future works. It also provide the limitations of the research.

REFERENCES

Aitkin, M., and Rubin, D. B. (1985). Estimation and hypothesis testing in finite mixture models. *Journal of the Royal Statistical Society: Series B (Methodological), 47*(1), 67-75.

Au, E. Y. L., and Goonetilleke, R. S. (2007). A qualitative study on the comfort and fit of ladies' dress shoes. *Applied Ergonomics, 38*(6), 687-696.

Awani, A. (2013). Shoe-making industry churns out RM1 billion in sales. *Astro Awani*. Retrieved 15 February 2016, from http://english.astroawani.com/business-news/shoe-making-industry-churns-out-rm1-billion-sales-11014#

Baba, K. (1974). Foot Measurement For Shoe Construction With Reference to The Relationship Between Foot Length, Foot Breadth, and Ball Girth *Journal of Human Ergology, 3*(2), 149-156.

Bari, S. B., Othman, M., and Mohd Salleh, N. (2010). Foot anthropometry for shoe design among preschool children in Malaysia. *Pertanika Journal of Social Sciences & Humanities, 18*(1), 69-79.

Bari, S. B., Salleh, N. M., Sulaiman, N., and Othman, M. (2015). Development of clothing size for pre school children based on anthropometric measurements *Australian Journal of Sustainable Business and Society, 1*(2).

Bataller, A., Alcántara, E., González, J., Garcia, A., and Alemany, S. (2001). *Morphological grouping of Spanish feet using clustering techniques.* Paper presented at the Proceedings of Fifth Symposium on Footwear Biomechanics, 12-13.

Baudry, J.-P., Raftery, A. E., Celeux, G., Lo, K., and Gottardo, R. (2010). Combining Mixture Components for Clustering. *Journal of Computational and Graphical Statistics, 19*(2), 332-353.

Beazley, A. (1998). Size and fit: Formulation of body measurement tables and sizing systems—Part 2. *Journal of Fashion Marketing and Management: An International Journal*.

Branthwaite, H., Chockalingam, N., and Greenhalgh, A. (2013a). The effect of shoe toe box shape and volume on forefoot interdigital and plantar pressures in healthy females. *J Foot Ankle Res, 6*(1), 28.

Branthwaite, H., Chockalingam, N., Grogan, S., and Jones, M. (2013b). Footwear choices made by young women and their potential impact on foot health. *Journal of health psychology, 18*(11), 1422-1431.

Chavent, M., Kuentz, V., Liquet, B. ı., and Saracco, L. (2011). ClustOfVar: an R package for the clustering of variables. *arXiv preprint arXiv:1112.0295*.

Cheng, F.-T., and Perng, D.-B. (2000). A systematic approach for developing a foot size information system for shoe last design. *International Journal of Industrial Ergonomics, 25*(2), 171-185.

Choi, Y.-L. (2012). Research on Growing Progress of Foot Size to Improve Appellation of Shoe Sizes. *Fashion & Textile Research Journal, 14*(3), 448-453.

Chua, Y. P., Tan, W. J., Yahya, T., and Saw, A. (2013). Prevalence of nontraumatic foot pain among urban young working women and its contributing factors. *Singapore Med J, 54*(11), 630-633.

Crawley, M. J. (2012). *The R book*: John Wiley & Sons.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1-38.

Everitt, B. (1993). *Cluster analysis* (3rd ed.): Edward Arnold and Halsted Press.

Everitt, B. S. (1981). *Finite mixture distributions*: Wiley Online Library.

Everitt, B. S. (2006). *An R and S-PLUS® companion to multivariate analysis*: Springer Science & Business Media.

Feng, Z. D., and McCulloch, C. E. (1996). Using bootstrap likelihood ratios in finite mixture models. *Journal of the Royal Statistical Society: Series B (Methodological), 58*(3), 609-617.

Fraley, C., and Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *The computer journal, 41*(8), 578-588.

Fraley, C., and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association, 97*(458), 611-631.

Fraley, C., and Raftery, A. E. (2007). Model-based methods of classification: using the mclust software in chemometrics. *Journal of Statistical Software, 18*(6), 1-13.

Gong, T., Fei, R., Lai, J., and Liang, G. (2010). *Foot shape analysis of adult male in the China.* Paper presented at the International Conference on Information Computing and Applications, 1-7.

Goonetilleke, R. S. (1999). Footwear cushioning: relating objective and subjective measurements. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(2), 241-256.

Goos, P., and Meintrup, D. (2016). *Statistics with JMP: Hypothesis Tests, ANOVA and Regression*: John Wiley & Sons.

Gupta, D., and Gangadhar, B. (2004). A statistical model for developing body size charts for garments. *International Journal of Clothing Science and Technology, 16*(5), 458-469.

Hamzah, M. F. M., Kasim, R. M., Yunus, A., Rijal, O. M., and Noor, N. M. (2017). *Detection of Interstitial Lung Disease using correlation and regression methods on texture measure.* Paper presented at the Imaging, Vision & Pattern Recognition (icIVPR), 2017 IEEE International Conference, 1-4.

Harrison, S., Cochrane, L., Abboud, R., and Leese, G. (2007). Do patients with diabetes wear shoes of the correct size? *International journal of clinical practice, 61*(11), 1900-1904.

Hawes, M. R., Sovak, D., Miyashita, M., Kang, S.-j., Yoshihuku, Y., and Tanaka, S. (1994). Ethnic differences in forefoot shape and the determination of shoe comfort. *Ergonomics*, *37*(1), 187-196.

Hawkins, D. M. (1976). Point estimation of the parameters of piecewise regression models. *Applied Statistics*, 51-57.

Hignett, S., and Masud, T. (2006). A review of environmental hazards associated with in-patient falls. *Ergonomics, 49*(5-6), 605-616.

Hill, M., Naemi, R., Branthwaite, H., and Chockalingam, N. (2017). The relationship between arch height and foot length: Implications for size grading. *Applied ergonomics, 59*, 243-250.

Hinojo-Pérez, J. J., Davia-Aracil, M., Jimeno-Morenilla, A., Sánchez-Romero, J. L., and Salas, F. (2016). Automation of the shoe last grading process according to international sizing systems. *The International Journal of Advanced Manufacturing Technology, 85*(1-4), 455-467.

Hisham, S., Mamat, C. R., and Ibrahim, M. A. (2012a). Multivariate statistical analysis for race variation from foot anthropometry in the Malaysian population. *Australian Journal of Forensic Sciences, 44*(3), 285-293.

Hisham, S., Mamat, C. R., and Ibrahim, M. A. (2012b). Regression analysis for stature estimation from foot anthropometry in Malaysian Chinese. *Australian Journal of Forensic Sciences, 44*(4), 333-341.

Hogg, R. V., Tanis, E. A., and Zimmerman, D. L. (1977). *Probability and statistical inference* (Vol. 993): Macmillan New York.

Hong, Y., Wang, L., Xu, D. Q., and Li, J. X. (2011). Gender differences in foot shape: a study of Chinese young adults. *Sports biomechanics, 10*(02), 85-97.

I-Ware Laboratory Co., L. (2017). INFOOT Measuring points and Items. Retrieved 20 December, 2017, from http://www.iwl.jp/content/mark_dimension.html

Ibrahim, R., and Khedifb, L. Y. B. (2004). Modeling 3D Surface of the Footprint-Malaysian Women Perspective.

Insights, M. (2014). Malaysia's Shoe Industry Steps Up.

Johnson, R. A., and Wichern, D. W. (2014). *Applied multivariate statistical analysis*: Pearson Education Limited Essex.

Kalebota, N., Szirovicza, L., Tucaković-Mujagić, L., and Drenovac, M. (2003). The structure of body measurements for the determination of shoe sizing for young Croatian men. *Collegium antropologicum, 27*(2), 635-643.

Kasahara, H., and Shimotsu, K. (2012). Testing the number of components in finite mixture models.

Kinanthropometry, I. S. S. A. i. (2010). ISO 20685: 2010, *ISAK*.

Klingenberg, C. P. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Development genes and evolution, 226*(3), 113-137.

Kong, P. W., Lim, C. Y., Ding, R., and Sterzing, T. (2015). Subjective evaluation of running footwear depends on country and assessment method: a bi-national study. *Ergonomics, 58*(9), 1589-1604.

Kouchi, M. (1998). Foot dimensions and foot shape: Differences due to growth, generation and ethnic origin. *Anthropological Science, 106*, 161-188.

Kouchi, M., and Mochimaru, M. (2011). Errors in landmarking and the evaluation of the accuracy of traditional and 3D anthropometry. *Applied ergonomics, 42*(3), 518-527.

Kouchi, M., and Yamazaki, N. (1992). Allometry of the foot and the shoe last. *Journal of the Anthropological Society of Nippon, 100*(1), 101-118.

Krauss, I., Grau, S., Mauch, M., Maiwald, C., and Horstmann, T. (2008). Sex-related differences in foot shape. *Ergonomics, 51*(11), 1693-1709.

Krauss, I., Valiant, G., Horstmann, T., and Grau, S. (2010). Comparison of Female Foot Morphology and Last Design in Athletic Footwear—Are Men's Lasts Appropriate for Women? *Research in Sports Medicine, 18*(2), 140-156.

Kwon, O., Jung, K., You, H., and Kim, H.-E. (2009). Determination of key dimensions for a glove sizing system by analyzing the relationships between hand dimensions. *Applied Ergonomics, 40*(4), 762-766.

Lee, Y.-C., Chao, W.-Y., and Wang, M.-J. (2012). *Foot shape classification using 3D scanning data.* Paper presented at the Network of Ergonomics Societies Conference (SEANES), 2012 Southeast Asian, 1-6.

Lee, Y.-C., KOUCHI, M., MOCHIMARU, M., and WANG, M.-J. (2015). Comparing 3D foot shape models between Taiwanese and Japanese females. *Journal of human ergology, 44*(1), 11-20.

Lee, Y.-C., Lin, G., and Wang, M.-J. J. (2014). Comparing 3D foot scanning with conventional measurement methods. *Journal of foot and ankle research, 7*(1), 44.

Lee, Y.-C., and Wang, M.-J. (2015). Taiwanese adult foot shape classification using 3D scanning data. *Ergonomics, 58*(3), 513-523.

Li, J. X., Xu, D. Q., Wang, A., and Shao, K. (2005). *Analysis of foot shape measures from Chinese male and female adults.* Paper presented at the Proceedings of the 7th Symposium on Footwear Biomechanics, 27-29.

Lin, C.-L., Wang, M.-J. J., and Drury, C. G. (2007). Biomechanical, physiological and psychophysical evaluations of clean room boots. *Ergonomics, 50*(4), 481-496.

Luximon, A. (2001). *Foot shape evaluation for footwear fitting.*

Luximon, A., Zhang, Y., Luximon, Y., and Xiao, M. (2012). Sizing and grading for wearable products. *Computer-Aided Design, 44*(1), 77-84.

Maindonald, J., and Braun, J. (2006). *Data analysis and graphics using R: an example-based approach* (Vol. 10): Cambridge University Press.

Manna, I., Pradhan, D., Ghosh, S., Kumar Kar, S., and Dhara, P. (2001). A comparative study of foot dimension between adult male and female and evaluation of foot hazards due to using of footwear. *Journal of physiological anthropology and applied human science, 20*(4), 241-246.

MATRADE. (2015). Positive Outlook for Malaysian Footwear and Accessories Exports in Europe Retrieved 15 February, 2016, from http://www.matrade.gov.my/en/142-press-releases/press-releases-2015/3245-positive-outlook-for-malaysian-footwear-and-accessories-exports-in-europe

Mauch, M., Grau, S., Krauss, I., Maiwald, C., and Horstmann, T. (2009). A new approach to children's footwear based on foot type classification. *Ergonomics, 52*(8), 999-1008.

McLachlan, G., and Peel, D. (2004). *Finite mixture models*: John Wiley & Sons.

McLachlan, G. J. (1987). On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 36*(3), 318-324.

McZgee, V. E., and Carleton, W. T. (1970). Piecewise regression. *Journal of the American Statistical Association, 65*(331), 1109-1124.

Miller, J. E., Nigg, B. M., Liu, W., Stefanyshyn, D. J., and Nurse, M. A. (2000). Influence of foot, leg and shoe characteristics on subjective comfort. *Foot & Ankle International, 21*(9), 759-767.

Mishra, M. K., Bruniaux, P., Crepin, D., and Hamad, B. (2017). STUDYING FOOT ANTHROPOMETRIC MEASURES USING 3D SCANNING METHOD. *Age (years), 32*(9.3), 36.38.

Mishra, P., Singh, U., Pandey, C. M., Mishra, P., and Pandey, G. (2019). Application of student's t-test, analysis of variance, and covariance. *Annals of Cardiac Anaesthesia, 22*(4), 407.

Moorthy, T. N., and Khan, H. B. M. A. (2013). Estimation of stature from foot outline (3D footprint) measurements in Ibans of East Malaysia by regression analysis. *International Journal of Biomedical and Advance Research, 4*(12), 889-895.

Moorthy, T. N., Ling, A. Y., Sarippudin, S. A., and Nik Hassan, N. F. (2014). Estimation of stature from footprint and foot outline measurements in Malaysian Chinese. *Australian Journal of Forensic Sciences, 46*(2), 136-159.

Moorthy, T. N., MohdKhalil, N. A. F. B., and Khan, H. M. A. (2013). Stature estimation based on footprint measurements of Malays in peninsular Malaysia by regression analysis. *International Journal of Biomedical and Advance Research, 4*(10), 683-689.

Most, R. S., and Sinnock, P. (1983). The epidemiology of lower extremity amputations in diabetic individuals. *Diabetes care, 6*(1), 87-91.

Muggeo, V. M. (2003). Estimating regression models with unknown break-points. *Statistics in medicine, 22*(19), 3055-3071.

Mündermann, A., Nigg, B. M., Stefanyshyn, D. J., and Humble, R. N. (2002). Development of a reliable method to assess footwear comfort during running. *Gait & posture, 16*(1), 38-45.

Nácher, B., Alemany, S., González, J., Alcántara, E., García-Hernández, J., Heras, S., et al. (2006). *A footwear fit classification model based on anthropometric data.* Paper presented at the Proceedings of the 8th annual digital human modelling for design and engineering symposium: 4-6th July 2006 Lyon.

Nataraja Moorthy, T., Mazidah, K., Hadzri, M., and Jayaprakash, P. (2011). Estimation of stature based on foot length of Malays in Malaysia. *Aust J Forensic Sci, 43*, 13-26.

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling, 14*(4), 535-569.

Oliveira-Brochado, A., and Martins, F. V. (2005). *Assessing the number of components in mixture models: a review*: Universidade do Porto, Faculdade de Economia do Portoo. Document Number)

Perini, T. A., Oliveira, G. L. d., Ornellas, J. d. S., and Oliveira, F. P. d. (2005). Technical error of measurement in anthropometry. *Revista Brasileira de Medicina do Esporte, 11*(1), 81-85.

Rawlings, J. O., Pantula, S. G., and Dickey, D. A. (2001). *Applied regression analysis: a research tool*: Springer Science & Business Media.

Reddy, P., Vaid, M., and Child, D. (1989). Diabetes and incorrectly fitting shoes. *Practical Diabetes, 6*(1), 16-16.

Rodrigo, A. S., Goonetilleke, R. S., and Witana, C. P. (2012). Model based foot shape classification using 2D foot outlines. *Computer-Aided Design, 44*(1), 48-55.

Romesburg, C. (2004). *Cluster analysis for researchers*: Lulu. com.

Roodbandi, A. S. J., Naderi, H., Hashenmi-Nejad, N., Choobineh, A., Baneshi, M. R., and Feyzi, V. (2017). Technical Report on the Modification of 3-Dimensional Non-contact Human Body Laser Scanner for the Measurement of

Anthropometric Dimensions: Verification of its Accuracy and Precision. *Journal of lasers in medical sciences, 8*(1), 22.

Rossi, W., and Tennant, R. (2000). Professional Shoe Fitting (New York: Pedorthic Footwear Association with acknowledgement to the National Shoe Retailers Association).

Sall, J., Lehman, A., Stephens, M. L., and Creighton, L. (2012). *JMP start statistics: a guide to statistics and data analysis using JMP*: Sas Institute.

Schroeder, L. D., Sjoquist, D. L., and Stephan, P. E. (1986). *Understanding regression analysis: An introductory guide* (Vol. 57): Sage.

Shah, S. A. (2018). Time for local shoemakers to innovate. *The Malaysian Reserve*. Retrieved 28 August 2019, from https://themalaysianreserve.com/2018/06/11/time-for-local-shoemakers-to-innovate/

Shariff, S. M., Bong, Y. B., Shariff, A. A., Kouchi, M., Kimura, K., Merican, A. F., et al. (2014). Development of a Foot Sizing System for Malaysian Women.

Shirkhorshidi, A. S., Aghabozorgi, S., and Wah, T. Y. (2015). A comparison study on similarity and dissimilarity measures in clustering continuous data. *PloS one, 10*(12).

Shugaba, A., Shinku, F., Gambo, I., Mohammed, M., Uzokwe, C., Damilola, E., et al. (2013). Relationship Between Foot Length, Foot Breadth, Ball Girth, Height and Weight of School Children Aged 3-5 Years Old.

Slater, K. (1986). Discussion paper the assessment of comfort. *Journal of the textile institute, 77*(3), 157-171.

Snedecor, G. W., and Cochran, W. G. (1980). *Statistical methods*: Iowa State University Press.

Standard, I. O. (2019). ISO 9407:2019, *Shoe sizes - Mondopoint system of sizing and marking*.

Standards, J. I. (1998). JIS S 5037:1998 *Sizing system for shoes*. Japan.

Standards, T. I. (2000). CNS 4800-S1093 :2000, *Sizing System for Shoes.* Taiwan.

Titterington, D. M., Smith, A. F. M., and Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*: Wiley.

Toms, J. D., and Lesperance, M. L. (2003). Piecewise regression: a tool for identifying ecological thresholds. *Ecology, 84*(8), 2034-2041.

Tu, H.-H. (2014). Foot volume estimation formula in healthy adults. *International Journal of Industrial Ergonomics, 44*(1), 92-98.

UNE, E. S. (1998). UNE 59850:1998 *Footwear - Designation Of Sizes - Main Characteristics Of The Continental System*. Spain, Madrid.

Vahidipour, S. M., Mirzaei, A., and Rahmati, M. (2014). Comparing weighted combination of hierarchical clustering based on Cophenetic measure. *Intelligent data analysis, 18*(4), 547-559.

Wang, C.-S. (2010). An analysis and evaluation of fitness for shoe lasts and human feet. *Computers in Industry, 61*(6), 532-540.

Watanabe, M., and Yamaguchi, K. (2003). *The EM algorithm and related statistical models*: CRC Press.

Waterson, P., and Sell, R. (2006). Recurrent themes and developments in the history of the Ergonomics Society. *Ergonomics, 49*(8), 743-799.

Wild, D. J. (2005). MINITAB release 14: ACS Publications.

Witana, C. P., Feng, J., and Goonetilleke, R. S. (2004). Dimensional differences for evaluating the quality of footwear fit. *Ergonomics, 47*(12), 1301-1317.

Witana, C. P., Xiong, S., Zhao, J., and Goonetilleke, R. S. (2006). Foot measurements from three-dimensional scans: A comparison and evaluation of different methods. *International Journal of Industrial Ergonomics, 36*(9), 789-807.

Xiong, S., Goonetilleke, R. S., Witana, C. P., and Lee Au, E. Y. (2008). Modelling foot height and foot shape-related dimensions. *Ergonomics, 51*(8), 1272-1289.

Yung-Hui, L., and Wei-Hsien, H. (2005). Effects of shoe inserts and heel height on foot pressure, impact force, and perceived comfort during walking. *Applied ergonomics, 36*(3), 355-362.

Zakaria, N. (2011). Sizing system for functional clothing-Uniforms for school children. *Indian Journal of Fibre & Textile Research, 36*(4), 348-357.

Zamarashkin, K., and Zamarashkin, N. (2009). Reference coordinate system in design of shoe lasts. *Footwear Science, 1*(sup1), 8-10.

Zhang, D., and Lu, G. (2004). Review of shape representation and description techniques. *Pattern recognition, 37*(1), 1-19.

Zhang, L., Drury, C. G., and Woolley, S. M. (1991). Constrained standing: evaluating the foot/floor interface. *Ergonomics, 34*(2), 175-192.

Zhang, L., Drury, C. G., and Woolley, S. M. (1991). Constrained standing: evaluating the foot/floor interface. Ergonomics, 34(2), 175-192.

Appendix A Comparison of CPCC Index of Different Distance Measures for 3 Foot measurements and Its Combination by Different Linkage Method.

(a) Linkage: Average — Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.7492	0.7492	0.7492	0.7385
FB	0.7756	0.7756	0.7756	0.7750
BG	0.7872	0.7872	0.7872	0.7752
FL and FB	0.7282	0.7141	0.7161	0.7156
FL and BG	0.7248	0.7190	0.7243	0.6773
FB and BG	0.7929	0.7874	0.7808	0.7807
FL, FB and BG	0.7411	0.7422	0.7237	0.7555

(b) Linkage: Single — Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.6966	0.6966	0.6966	0.6730
FB	0.6780	0.6780	0.6780	0.6550
BG	0.6988	0.6988	0.6988	0.6820
FL and FB	0.6018	0.6143	0.5903	0.5978
FL and BG	0.5828	0.5831	0.6295	0.5658
FB and BG	0.5696	0.5343	0.5991	0.3639
FL, FB and BG	0.5270	0.5171	0.5833	0.4182

(c) Linkage: Complete — Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.6472	0.6472	0.6472	0.6441
FB	0.6729	0.6729	0.6729	0.7016
BG	0.7567	0.7567	0.7567	0.7284
FL and FB	0.5850	0.6119	0.6244	0.6541
FL and BG	0.5649	0.5783	0.6341	0.5977
FB and BG	0.6188	0.6259	0.7222	0.7476
FL, FB and BG	0.6419	0.6099	0.6099	0.5665

(d) Linkage: Centroid — Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.7548	0.7548	0.7548	0.7454
FB	0.7719	0.7719	0.7719	0.7371
BG	0.7718	0.7718	0.7718	0.7758
FL and FB	0.7206	0.7092	0.7108	0.7106
FL and BG	0.7393	0.7209	0.6934	0.7091
FB and BG	0.7907	0.7932	0.7879	0.7587
FL, FB and BG	0.7488	0.7499	0.7283	0.7484

(e) Linkage: Median Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.6537	0.6537	0.6537	0.7028
FB	0.7059	0.7059	0.7059	0.7452
BG	0.7636	0.7636	0.7636	0.6480
FL and FB	0.4936	0.5423	0.5761	0.5065
FL and BG	0.5263	0.6361	0.6223	0.5780
FB and BG	0.7612	0.6601	0.7135	0.7481
FL, FB and BG	0.5668	0.5893	0.5705	0.6174

(f) Linkage: Ward Distance measure

Foot shape combination	Euclidean	Manhattan	Maximum	Canberra
FL	0.6330	0.6330	0.6330	0.6291
FB	0.6292	0.6292	0.6292	0.6000
BG	0.6126	0.6126	0.6126	0.6251
FL and FB	0.5963	0.6012	0.5358	0.5437
FL and BG	0.5964	0.5639	0.5252	0.5706
FB and BG	0.6101	0.5989	0.6002	0.6940
FL, FB and BG	0.5745	0.5802	0.4893	0.5902

* the yellow box indicates the highest value of CPCC index for a fix foot measurement combination.

Appendix B Comparison of CPCC Index of Different Distance Measures for 5 Foot measurements and Its Combination by Different Linkage Method

	Average				Single			
Shape combination	Euclidean	Manhattan	Maximum	Canberra	Euclidean	Manhattan	Maximum	Canberra
FL	0.749	0.749	0.749	0.739	0.697	0.697	0.697	0.673
FB	0.776	0.776	0.776	0.775	0.678	0.678	0.678	0.655
BG	0.787	0.787	0.787	0.775	0.699	0.699	0.699	0.682
IL	0.788	0.788	0.788	0.788	0.704	0.704	0.704	0.708
FIL	0.782	0.782	0.782	0.777	0.674	0.674	0.674	0.659
FL, FB	0.728	0.714	0.716	0.716	0.602	0.614	0.590	0.598
FL, BG	0.725	0.719	0.724	0.677	0.583	0.583	0.629	0.566
FL, IL	0.763	0.764	0.762	0.753	0.707	0.696	0.697	0.687
FL, FIL	0.769	0.690	0.776	0.673	0.524	0.502	0.592	0.497
FB, BG	0.793	0.787	0.781	0.781	0.570	0.534	0.599	0.364
FB, IL	0.698	0.681	0.707	0.660	0.671	0.660	0.665	0.568
FB, FIL	0.736	0.759	0.750	0.722	0.710	0.637	0.737	0.569
BG, IL	0.763	0.748	0.733	0.706	0.678	0.655	0.663	0.614
BG, FIL	0.787	0.786	0.778	0.752	0.630	0.653	0.570	0.621
IL, FIL	0.788	0.764	0.803	0.754	0.622	0.539	0.697	0.543
FL, FB, BG	0.741	0.742	0.724	0.755	0.527	0.517	0.583	0.418
FL, FB, IL	0.719	0.723	0.699	0.684	0.584	0.579	0.569	0.555
FL, FB, FIL	0.684	0.665	0.765	0.648	0.528	0.480	0.587	0.497
FL, BG, IL	0.714	0.671	0.715	0.692	0.635	0.613	0.669	0.584
FL, BG, FIL	0.753	0.692	0.743	0.711	0.585	0.548	0.606	0.534
FL, FIL,IL	0.731	0.721	0.766	0.661	0.528	0.475	0.622	0.461
FB, BG, IL	0.760	0.766	0.734	0.752	0.603	0.547	0.644	0.470
FB, BG, FIL	0.788	0.781	0.757	0.767	0.594	0.559	0.558	0.476
FB, IL, FIL	0.748	0.745	0.787	0.711	0.628	0.537	0.690	0.517
BG, IL, FIL	0.752	0.727	0.756	0.713	0.636	0.604	0.648	0.545
FL, FB, BG, IL	0.716	0.683	0.718	0.682	0.551	0.524	0.646	0.467
FL, FB, BG, FIL	0.738	0.732	0.742	0.731	0.561	0.508	0.599	0.465
FL, BG,IL, FIL	0.670	0.681	0.752	0.678	0.618	0.554	0.635	0.521
FB, BG,IL, FIL	0.763	0.722	0.755	0.703	0.623	0.555	0.648	0.488
FL, FB,IL,FIL	0.717	0.716	0.769	0.686	0.545	0.496	0.627	0.508
FL,FB,BG,IL,FIL	0.709	0.705	0.755	0.677	0.590	0.524	0.631	0.468
	Complete				Centroid			
Shape combination	Euclidean	Manhattan	Maximum	Canberra	Euclidean	Manhattan	Maximum	Canberra
FL	0.647	0.647	0.647	0.644	0.755	0.755	0.755	0.745
FB	0.673	0.673	0.673	0.702	0.772	0.772	0.772	0.737
BG	0.757	0.757	0.757	0.728	0.772	0.772	0.772	0.776
IL	0.610	0.610	0.610	0.622	0.785	0.785	0.785	0.781
FIL	0.588	0.588	0.588	0.553	0.803	0.803	0.803	0.776
FL, FB	0.585	0.612	0.624	0.654	0.721	0.709	0.711	0.711
FL, BG	0.565	0.578	0.634	0.598	0.739	0.721	0.693	0.709
FL, IL	0.635	0.626	0.713	0.642	0.761	0.726	0.759	0.747
FL, FIL	0.710	0.707	0.650	0.590	0.751	0.679	0.759	0.718
FB, BG	0.619	0.626	0.722	0.748	0.791	0.793	0.788	0.759
FB, IL	0.570	0.509	0.540	0.584	0.729	0.701	0.713	0.681
FB, FIL	0.578	0.605	0.650	0.653	0.769	0.732	0.741	0.699
BG, IL	0.623	0.636	0.729	0.556	0.757	0.730	0.753	0.703
BG, FIL	0.715	0.605	0.561	0.648	0.772	0.767	0.760	0.729
IL, FIL	0.568	0.655	0.527	0.597	0.771	0.747	0.775	0.734
FL, FB, BG	0.642	0.610	0.610	0.566	0.749	0.750	0.728	0.748
FL, FB, IL	0.591	0.585	0.590	0.567	0.746	0.716	0.708	0.708
FL, FB, FIL	0.629	0.579	0.533	0.571	0.716	0.721	0.718	0.700
FL, BG, IL	0.545	0.674	0.593	0.571	0.721	0.685	0.717	0.702
FL, BG, FIL	0.698	0.600	0.583	0.671	0.712	0.690	0.710	0.684
FL, FIL,IL	0.635	0.598	0.689	0.592	0.737	0.724	0.662	0.736
FB, BG, IL	0.619	0.614	0.663	0.629	0.764	0.747	0.755	0.748

FB, BG, FIL	0.618	0.716	0.583	0.662	0.781	0.771	0.758	0.750
FB, IL, FIL	0.663	0.594	0.621	0.641	0.739	0.723	0.751	0.696
BG, IL, FIL	0.647	0.662	0.607	0.552	0.712	0.682	0.745	0.703
FL, FB, BG, IL	0.610	0.591	0.659	0.615	0.726	0.699	0.727	0.690
FL, FB, BG, FIL	0.695	0.584	0.682	0.589	0.731	0.699	0.729	0.688
FL, BG,IL, FIL	0.663	0.562	0.686	0.533	0.679	0.691	0.697	0.706
FB, BG,IL, FIL	0.695	0.640	0.573	0.538	0.715	0.709	0.744	0.703
FL, FB,IL,FIL	0.617	0.597	0.679	0.546	0.707	0.744	0.710	0.713
FL,FB,BG,IL,FIL	0.608	0.549	0.681	0.572	0.683	0.674	0.722	0.678

	Median				Ward			
Shape combination	Euclidean	Manhattan	Maximum	Canberra	Euclidean	Manhattan	Maximum	Canberra
FL	0.654	0.654	0.654	0.654	0.654	0.654	0.654	0.654
FB	0.706	0.706	0.706	0.706	0.706	0.706	0.706	0.706
BG	0.764	0.764	0.764	0.764	0.764	0.764	0.764	0.764
IL	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680
FIL	0.759	0.759	0.759	0.759	0.759	0.759	0.759	0.759
FL, FB	0.494	0.542	0.494	0.542	0.494	0.542	0.494	0.542
FL, BG	0.526	0.636	0.526	0.636	0.526	0.636	0.526	0.636
FL, IL	0.715	0.610	0.715	0.610	0.715	0.610	0.715	0.610
FL, FIL	0.676	0.591	0.676	0.591	0.676	0.591	0.676	0.591
FB, BG	0.761	0.660	0.761	0.660	0.761	0.660	0.761	0.660
FB, IL	0.556	0.462	0.556	0.462	0.556	0.462	0.556	0.462
FB, FIL	0.639	0.590	0.639	0.590	0.639	0.590	0.639	0.590
BG, IL	0.432	0.504	0.432	0.504	0.432	0.504	0.432	0.504
BG, FIL	0.652	0.533	0.652	0.533	0.652	0.533	0.652	0.533
IL, FIL	0.652	0.637	0.652	0.637	0.652	0.637	0.652	0.637
FL, FB, BG	0.567	0.589	0.567	0.589	0.567	0.589	0.567	0.589
FL, FB, IL	0.672	0.573	0.672	0.573	0.672	0.573	0.672	0.573
FL, FB, FIL	0.649	0.591	0.649	0.591	0.649	0.591	0.649	0.591
FL, BG, IL	0.692	0.611	0.692	0.611	0.692	0.611	0.692	0.611
FL, BG, FIL	0.569	0.635	0.569	0.635	0.569	0.635	0.569	0.635
FL, FIL,IL	0.518	0.611	0.518	0.611	0.518	0.611	0.518	0.611
FB, BG, IL	0.578	0.601	0.578	0.601	0.578	0.601	0.578	0.601
FB, BG, FIL	0.631	0.734	0.631	0.734	0.631	0.734	0.631	0.734
FB, IL, FIL	0.685	0.510	0.685	0.510	0.685	0.510	0.685	0.510
BG, IL, FIL	0.606	0.519	0.606	0.519	0.606	0.519	0.606	0.519
FL, FB, BG, IL	0.567	0.482	0.567	0.482	0.567	0.482	0.567	0.482
FL, FB, BG, FIL	0.703	0.595	0.703	0.595	0.703	0.595	0.703	0.595
FL, BG,IL, FIL	0.515	0.650	0.515	0.650	0.515	0.650	0.515	0.650
FB, BG,IL, FIL	0.748	0.579	0.748	0.579	0.748	0.579	0.748	0.579
FL, FB,IL,FIL	0.624	0.562	0.624	0.562	0.624	0.562	0.624	0.562
FL,FB,BG,IL,FIL	0.537	0.612	0.537	0.612	0.537	0.612	0.537	0.612

* the green box indicates the highest value of CPCC index for all distance measure of six different linkage method of a certain foot measurement combination.

Appendix C R-Command for Statistical Approach

Cluster of Features to Select Foot Measurements

```
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan data\\Left foot_PG\\All features left.txt",header=T)
cor(LF1)
d.corr <- as.dist(1-cor(LF1)^2)
d.corr
#plot dendrogram
hc <- hclust(d.corr, method='average')
plot(as.dendrogram(hc))
#calculate CPCC
d2 <- cophenetic(hc)
cor (d.corr,d2)

cor(LF1)
d.corr <- as.dist(cor(LF1))
d.corr
#plot dendrogram
hc <- hclust(d.corr, method='ward')
plot(as.dendrogram(hc))
#calculate CPCC
d2 <- cophenetic(hc)
cor (d.corr,d2)
```

Outlier Detection From Boxplot, 2D plot and 3D plot

```
#boxplot with outlier label
library(car)
#For combine group with label

Boxplot(LFL ~ hc2, data=temp,
        col = "lightgray",
        main = "Foot Measuremet for Foot Length (FL)",
        xlab = "Group",
        ylab = "FL")
Boxplot(LFB ~ hc2, data=temp,
        col = "lightgray",
        main = "Foot Measuremet for Foot Breadth (FB)",
        xlab = "Group",
        ylab = "FB")
Boxplot(LBGC ~ hc2, data=temp,
        col = "lightgray",
        main = "Foot Measuremet for Ball Girth (BG)",
        xlab = "Group",
        ylab = "BG")
```

```
#To plot 2-D graph based on cluster result
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan data\\Left foot_PG\\left11.txt",header=T)
d1 <- dist(LF1,method =  "euclidean")
hc <- hclust(d1, method = "average")
hc2<-cutree(hc,k=2)
dframe = data.frame(LF1, hc2)
temp<-cbind(LF1,hc2)
plot(temp$LFL,temp$LFB,pch = 21, bg=c("red","blue")[unclass(temp$hc2)],
main = "Foot Breadth (FB) Vs Foot Length (FL)", xlab="FL", ylab="FB")
plot(temp$LFL,temp$LBGC,pch = 21, bg=c("red","blue")[unclass(temp$hc2)],
main = "Ball Girth (BG) Vs Foot Length (FL)", xlab="FL", ylab="BG")
plot(temp$LBGC,temp$LFB,pch = 21, bg=c("red","blue")[unclass(temp$hc2)],
main = "Foot Breadth (FB) Vs Ball Girth (BG)", xlab="BG", ylab="FB")

#To plot 3-D graph based on cluster result
library(rgl)
??plot3d
plot3d(temp,type= 's', col=temp$hc2,size=1)
```

<u>Clustering of Subject</u>

```
library(mixtools)
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan data\\Left foot_PG\\left11_3.txt",header=T)
d1 <- dist(LF1,method =  "euclidean")
hc <- hclust(d1, method = "average")
d2 <- cophenetic(hc)
cor (d1,d2)
hc1<-cutree(hc,k=2)
library(doBy)
dframe1 = data.frame(LF1, hc1)
summaryBy(LFL~hc1, data=dframe1, FUN=c(mean,sd))
summaryBy(LBGC~hc1, data=dframe1, FUN=c(mean,sd))
summaryBy(LFB~hc1, data=dframe1, FUN=c(mean,sd))
for(i in 4){dframe1[,i]<-factor(dframe1[,i])}Ba#assign hc1 as factor
summary (dframe1)
temp<-cbind(LF1,hc1)
str(temp)
cov(temp[hc1==1,1:3])
cov(temp[hc1==2,1:3])
LF1<-as.matrix(LF1)
a<-boot.comp(LF1, x = NULL, N = NULL, max.comp = 2, B = 100,
        sig = 0 ,mix.type = "mvnormalmix", hist = TRUE)
a
a$p.value
```

Multivariate Normal Mixture

```
lambda <- c(0.8625,0.1375)
mu <- list(c(232.15, 94.11, 225.77), c(247.53,107.64,257.82))
sigma <-
list(matrix(c(109.879,27.312,66.058,27.312,20.378,46.609,66.058,46.609,118.710),
3, 3),
matrix(c(134.837,5.218,16.031,5.218,7.643,15.267,16.031,15.267,40.687), 3, 3))
library( mixtools)
set.seed(100)
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan
data\\Left foot_PG\\left11_3.txt",header=T)
LFLLFBLBGC<-
mvnormalmixEM(LF1,lambda=lambda,mu=mu,sigma=sigma,arbvar = TRUE, k=2)
LFLLFBLBGC<- mvnormalmixEM(LF1,k=3)
summary (LFLLFBLBGC)
LFLLFBLBGC$mu
LFLLFBLBGC$sigma
LFLLFBLBGC$lambda
```

Piecewise Regression

```
##Cluster of features
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan
data\\Left foot_PG\\left all_3.txt",header=T)
##dendrogram by using corr as distance
cor(LF1)
d.corr <- as.dist(cor(LF1))
d.corr
#plot dendrogram
hc <- hclust(d.corr, method='ward')
plot(as.dendrogram(hc))
#calculate CPCC
d2 <- cophenetic(hc)
cor (d.corr,d2)

##piece wise for FB vs BG based on highest R2 value
LF1<-read.table("C:\\Faizal Hamzah\\Faizal Hamzah\\PhD\\Proposal\\Foot scan
data\\Left foot_PG\\left all_3.txt",header=T)
x<-LF1$BG
y<-LF1$FB
#plot normal simple regression
LR<-lm (y~x,data=LF1)
summary(LR)
plot (y~x, pch=16, cex=0.9, xlab="BG", ylab="FB")
abline(lm(y~x, data=LF1), col="red")

#sort BG
Break<-(sort(unique(x))[2:160])
```

```
Break
#working now yeay
mse<-numeric(131)
for (i in 1:131) {
model<-lm(y~(x<Break[i])*x + (x>=Break[i])*x)
mse[i]<-summary(model)[[6]] }
mse <- as.numeric(mse)
mse
plot(mse, pch=16)
#one possible breakpoint 232.5
piecewise <- lm(y ~ x*(x < 232.5) + x*(x > 232.5))
summary(piecewise)
```

Appendix D Mathematical Approach To Subdivide The Ellipsoid

Given $\underline{x} \sim N_p\left(\underline{\mu}, \Sigma\right)$, then we used

$$Pr\left[\left(\underline{x}-\underline{\mu}\right)^T \Sigma^{-1}(\underline{x}-\underline{\mu}) \leq C\right] \quad \text{(D. 1)}$$

where $C \sim \chi^2(p)$. If $y \leq C$ where $y \sim \chi^2(p)$, then ellipsoid can be represented as;

$$\left(\underline{x}-\underline{\mu}\right)^T \Sigma^{-1}\left(\underline{x}-\underline{\mu}\right) = C \quad \text{(D.2)}$$

Estimate of $\underline{\mu}$ is given by

$$\underline{\hat{\mu}} = \overline{\underline{x}} = \frac{1}{n}\left(\underline{x}_1 + \underline{x}_2 + \cdots + \underline{x}_n\right) \quad \text{(D. 3)}$$

where $\underline{x}_k$ represent the k-th subject (person),

and the estimate of $\sum$ is given by

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\underline{x}_i - \overline{\underline{x}}\right)\left(\underline{x}_i - \overline{\underline{x}}\right)^T \quad \text{(D. 4)}$$

In equation (D.1), (6.2), (6.3) and (6.4), $\underline{\mu}$ and $\sum$ from the k-th ellipsoid is replaced with $\underline{\mu}_k$ and $\sum_k$, $k = 1,2$. The value of n in (D.4) is replaced by n_k for the k-th ellipsoid, $k = 1,2$. Subdivide the ellipsoid into small 3-D subdivisions is equivalent to subdivide the respective sphere since it can be shown that,

$$N_p\left(\underline{\mu}, \Sigma\right) \rightarrow N_p\left(\underline{0}, I\right) \quad \text{(D. 5)}$$

using $\underline{v} = \underline{z} - w$ (D.6)

$$\underline{v} = \Lambda^{-\frac{1}{2}}\left(Q^T\left[\underline{x}-\underline{\mu}\right]\right) \quad \text{(D. 7)}$$

Estimates of Q and Λ is given from $\hat{\Sigma} = \hat{Q} \wedge \hat{Q}^T$ Result in (D.5) will be illustrated in the following 2 pages.

Suppose $\underline{x} = \left(x_1, x_2, \cdots, x_p\right)$ where the individual $\underline{x}$ is described by p-foot measurements (for example x_1 equal foot length). Let $\underline{x} \sim N_p\left(\underline{\mu}, \Sigma\right)$. Then the following standard theorems may be used.

Theorem D (a);

For $A^T \in \mathrm{M}_{qp}$ and $\underline{b} \in \mathbb{R}^q$, let $\underline{y} = A^T\underline{x} + \underline{b}$.

Then $\underline{y} \sim N_q\left(A^T\underline{x} + \underline{b}, A^T \sum A\right)$ $\quad$ (D.8)

Theorem D (b);

If $\underline{x} \sim N_p\left(\underline{\mu}, \sum\right)$ and $\sum$ is non-singular. Then,

$$y = \left(\underline{x} - \underline{\mu}\right)^T \Sigma^{-1} \left(\underline{x} - \underline{\mu}\right) \sim \chi^2(p) \quad \text{(D.9)}$$

Also, $y = c$ is an ellipsoid, say $c = \chi^2_{0.05}(p)$

Theorem D (c); Spectral decomposition of symmetric matrices

If A is $(n \times n)$ symmetric matrix with eigen value $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and eigen vector $(q_1, q_2, \dots, q_n)$.

When A$q_i = \lambda_i q_i$ and $q_i^T q_i = 1$ with $q_i^T q_j = 0 (i \neq j)$, then

$$\mathrm{A} = Q\Lambda Q^T \quad \text{(D.10)}$$

where $Q = \left[\underline{q}_1 \left|\underline{q}_2\right| \dots |\underline{q}_p\right]$ is an orthogonal $Q^T Q = 1$

and $\underline{\Lambda} = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$

ELLIPSOID TO UNIT SPHERE

Let $\underline{x} \sim N(\underline{\mu}, \sum)$, $\underline{\mu}$ is a set of mean for foot length, foot breadth and ball girth, and $\sum$ is a set of covariance for foot length, foot breadth and ball girth. We know that $\sum = Q\Lambda Q^T$ is an orthogonal.

Then,

$$y = Q^T\underline{x} \sim N_p\left(Q^T\underline{\mu}, Q^T \Sigma Q\right) \quad \text{(D.11)}$$

Since $\sum = Q\Lambda Q^T$, then

$$Q^T \Sigma Q = Q^T (Q\Lambda Q^T) Q$$

$$= \Lambda$$

$$Q^T \underline{x} \sim N_p\left(Q^T\underline{\mu}, \Lambda\right) \quad \text{(D.12)}$$

Let

$$\Lambda^{\frac{1}{2}} = \begin{pmatrix} \lambda_1^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & \lambda_2^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p^{\frac{1}{2}} \end{pmatrix}$$

$$\Lambda^{\frac{1}{2}}\,\Lambda^{\frac{1}{2}} = \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

Let $\underline{z} = \Lambda^{-\frac{1}{2}}\, Q^T\, \underline{x}$ (D.13)

$\therefore\ \underline{z} \sim N_p\left(\Lambda^{-\frac{1}{2}}\, Q^T\, \underline{\mu}\,, \mathrm{I}\right)$ (D.14)

Given that

$$\Lambda^{-\frac{1}{2}}\,\Lambda\,\Lambda^{-\frac{1}{2}} = \Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}\,\Lambda^{\frac{1}{2}}\,\Lambda^{-\frac{1}{2}}$$

$$\Lambda^{-\frac{1}{2}}\,\Lambda\,\Lambda^{-\frac{1}{2}} = I_p\, I_p = I_p$$

$\therefore\ \underline{z} \sim N_p\left(\underline{w}, \mathrm{I}\right)$ (D.15)

where $\underline{w} = \Lambda^{-\frac{1}{2}}\, Q^T\, \underline{\mu}$ (D.16)

Finally, let $\underline{v} = \underline{z} - \underline{w}$

$\therefore\ \underline{v} \sim N_p\left(\underline{0}, \mathrm{I}\right)$ (D.17)

Now, we slice up ellipsoid from (D.11) and convert back result in term of $\underline{x}$. Note that,

$\underline{v} = \underline{z} - \underline{w} = \Lambda^{-\frac{1}{2}}\, Q^T\left(\underline{x} - \underline{\mu}\right)$ (D.18)

APPLICATION OF UNIT SPHERE FOR SIZING SYSTEM

Let $\underline{v}^T = \left(v_1, v_2, \cdots, v_p\right)$ and $\underline{x}^T = \left(\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_p\right)$. We have $\underline{v} = \Lambda^{-\frac{1}{2}}\, Q^T(\underline{x} - \underline{\mu})$ from (D.13). To get back to $\underline{x}$ we use the following

$\underline{x} = (Q^T)^{-1}\Lambda^{\frac{1}{2}}\, \underline{v} + \underline{\mu}$ (D.19)

Q^T is nonsingular, so that Q^{-T} is an unique and there will be one to one correspond between $\underline{x}$ and $\underline{\mu}$. From theorem D (b) and result from (D.17)

$$y = \left(\underline{v} - \underline{0}\right)^T I^{-1}\left(\underline{v} - \underline{0}\right) \sim \chi^2(p) \qquad \text{(D.20)}$$

The quadratic term in (D.20) and result (D.20), and recall $p = 3$ in section D.1, therefore

$$v_1^2 + v_2^2 + v_3^2 = c \qquad \text{(D.21)}$$

where $c = \chi^2_{0.95}(3)$ from standard Chi-squared table.

The values of v_1, v_2 and v_3 are data dependent and once they are selected, equation D.19 give the foot measurement.

LIST OF PUBLICATIONS

Hamzah, M. F. M., Rijal, O. M., Mokhtar, N. A., and Noor, N. M. (2017a). Piecewise regression for determining Malaysian woman foot shape. Paper presented at the 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 379-384.

Hamzah, M. F. M., Rijal, O. M., and Noor, N. M. (2017b). Validation of clusters for M-foot shape measurement: Malaysian women foot shape. Paper presented at the Region 10 Conference, TENCON 2017-2017 IEEE, 906-911.

Rijal, O. M., Hamzah, M. F. M., Sankaraiah, S., and Noor, N. M. (2018). A three dimensional foot fourier descriptors model. ARPN Journal of Engineering and Applied Sciences, 13(13), 4042-4056.

Hamzah, M. F. M., Rijal, O. M., Kimura, K., and Noor, N. M. (2019). Malaysian Women Shoe Sizing System Using Multivariate Normal Probability Distribution. IEEE Access, 7, 142726-142737.