International Journal of Statistics and Applied Mathematics 2021; 6(5): 91-108

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(5): 91-108

www.mathsjournal.com
Received: 02-06-2021
Accepted: 07-07-2021

Unyime Patrick Udoudo
Department of Statistics, Akwa Ibom State Polytechnic, Ikot Osurua, Nigeria

Chisimkwuo John
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Nigeria

Kenny Ugochukwu Urama
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Nigeria

Nelson Theophilus Obijuru
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Nigeria

Corresponding Author:
Chisimkwuo John
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Nigeria

Discriminant analysis of anthropometric measurements: A unisex ready-made trousers Case study

Unyime Patrick Udoudo, Chisimkwuo John, Kenny Ugochukwu Urama and Nelson Theophilus Obijuru

Abstract
Fit to size is an important factor for consumers engaged into wearing unisex ready-made trousers. This study demonstrates how discriminant analysis can be used to improve the standard sizing system of ready-made trousers in Nigeria. To achieve this, the researcher measured 400 randomly selected students (200 male and 200 female) of Michael Okpara University of Agriculture, Umudike on seven anthropometric measurements; Waist, Hip, Length, Lap, Knee, Flap, and Base used in sewing trousers students. The data grouped into three (3) different sizes (small -s, large -l, and medium -m) on the basis of existing standard sizing measurement system. The assumptions of normality were established while that of equal covariance matrices was invalidated. Thus, the quadratic classification rule was employed per pair of the variables considered. On the overall, the analysis showed that about 3% of the sizes sampled were misclassified but when the derived classification function was evaluated it correctly classified about 97% of the sizes under study.

Keywords: Multivariate normality, discriminant function, Mahalonobis distance, T^2, Q-Q Plots

Introduction

Anthropometry is the science that measures the range of body size in a population. Many scholars agree on the need to measure human body dimensions in order to develop standards and solve variation in body size due to different reasons such as geographical location, nutrition, and ethnic group etc. Anthropometric data of a country are vital database for clothing design and other design applications. It is also an important parameter in population studies. Developing such a database is common in many other countries of the world. In Nigeria, however measuring human body dimensions have been given limited attention though Nigerian body dimensions are significantly different from others. Due to this, Nigerian (African) people are being challenged with unfit products, machines, equipment, etc.

In the clothing/textile industry, designing in different size to address majority of the population is a common approach. However, in Nigeria, body measurement is mostly made by tailors to design cloths that fit individuals. Though custom made products have their own advantages; the people prefer to buy ready-made cloths for their cheap prices and accessibility in a short time.

Moreover, fit is one of the most important criteria for consumers engaged in the apparel buying decision. To get the best fit and size dimensions, manufacturers speed large amounts of money on sizing system. Every garment manufacturer has a target segment with certain demographic characteristics defining consumers′ profile. The best range of sizing can be a key to the success factor for manufacturers. To implement this, companies are using advanced technologies and strategies to devise sizing system and sizing categories (Doshi, 2006) [13].

Given that ready-made apparel depend on an accurate estimate of the distribution of body shapes and sizes within a target population, it becomes necessary for every country, even regions within countries to establish their own sizing systems based on the target population (Ashdown 2000; Simmons, Istook and Devarajan, 2004; Honey and Olds, 2007) [4, 12, 9].

Dissatisfaction with fit is one of the most frequently stated problems with garment purchase (Otieno, Harrow and Lea-Greenwood, 2005; Zwane and Magagula, 2006; Mastamet-mason, 2008) [31, 41, 26]. Consumers' dissatisfaction with apparels' fit has stimulated the development of assistive technologies in manufacturing, processing and retail environment such as the use of body scanners. Through body scan technology, body dimensions and shapes can easily and rapidly be extracted from a population and converted immediately into body form categories, size charts and patterns for garment production (Ashdown, 1998, Simmons and Istook, 2003; Ulrich, Anderson-Connell, and Wu, 2003; Ashdown, Loker, and Adelson, 2004; Fiore, Lee and Kunz, 2004) [3, 34, 37, 6, 15].

Due to costs and technical requirements, body scan technology would not be feasible in a less developed country. African developing countries such as Nigeria also face similar apparel fit problems, but sizing issues are often overlooked or regarded as unimportant issues, finally given rise to non-standardized size ranges that do not conform to the recommendations given by standard bodies (Chun-Yoon and Jasper, 1995; Faust, Carrier and Baptiste, 2006) [5, 14]. A lack of basic design technologies such as computer-aided design and pattern design systems, in most apparel industries are an indication of the ignorance about the importance of size and fit, which reveal a reluctance to respond to consumer demands (Mason, 1998) [25]. Imported new and second-hand clothes cannot address fit problems either because they were not manufactured for the Nigerian market.

Among others, unavailability of standard anthropometric data is one of the causes for hot producing ready-made cloths in the garment shops. Moreover, since mass production often bring economies of scale in product design through reduced setup time and stoppage. It is a desired condition for the sector development. These circumstances inculcate the development of anthropometric data to determine a minimum of different sizes (and the dimensions of each size) that will accommodate all users. Due to this, different countries have their own standard body dimension or size standards. Consequently, garment manufacturers produce different garments for male and female specifically for the intended people from the standard body measurement or size.

A good idea of one's shape is determined by measuring the circumference of the waist, hip and chest. Height is another important factor; two people who have the same height, hip and long arms and legs, others have long waist. These characteristics make a big difference in the kind of style that fit and flatten each individual.

Furthermore, since each outfit you wear creates a visual image that says something about you, this is why people are conscious of their own clothes and this is a challenge to those who design ready-made trousers that must be addressed accordingly. Thus, the major concern of this study was to attempt to separate and classify these anthropometric measurements or data into three categories or groups (sizes), small, medium and large using standard sizing system or set of standard body size.

Problem

In Nigeria, there are some people (male and female) who wear closely fitting clothing while others wear clothes which are oversize. However, there are those who wear clothes that fit will, while others wear clothes which are either too long or short for their height. McCormick, Kimuyu and Kinyanyui (2002) [29], found that most of the personnel in apparel industry are inadequately skilled to tackle fit issues and seldom employ modern technologies or even utilize dress firms for testing the fit of prototype apparel before even engaging fit models.

With the absence of representative sizing systems in Nigeria, it is possible that wrong styles and sizes based on estimates and not on the actual sizes and body shapes of male and female consumers in Nigeria contributes to fit and size problems. Consumers' ignorance of their own sizes and body shape leads to inappropriate apparel selection. It is also possible that consumers' fit preferences contribute to fit problems. If a consumer prefers loosely fitted garments on a thin body framework, the aesthetic appearance will be awkward and will appear ill-fitting.

However, it is a well known fact that some men have feminine figures and some women have masculine figures. This is why it is difficult to know the actual sizes of trouser for such people. Therefore, we need to distinguish them. Thus, clothing or fashion industries in Nigeria are faced with the problem of incorrect or no standard sizing system for producing unisex ready-made trousers. The question now is how we can set a correct or standard sizing system that can be used to design or produce unisex ready-made trousers for all the stakeholders in the business? In light of this, this study sought to separate and classify the male and female measurements into different sizes (small, medium and large) for sewing a unisex or universal ready-made trouser based on a standard sizing system.

Methodology

Method of Data Collection

The data for this study were collected from a primary source. A simple random of two hundred (200) students each for male and female were included in this study, making a total sample size of 400 (four hundred) students. The data were made up of anthropometric measurements used in sewing a pair of trousers such as waist, hip, length, lap, knee, flap and base. The data were further classified into three different sizes according to a set of standard sizing system (range of sizes) such as small size (28-30), medium size (32-34) and large size (36 and above). These sizes were arranged based on the waist measurement. All were measured using tailors' measuring tape and recorded in inches.

Method of Data Analysis

Since the study aims at predicting the suitable size (small, medium and large) of the anthropometric measurement based on the basis of their standard sizing system, the data were analyzed using discriminant analysis.

Discriminant Analysis

Discriminant Analysis is a multivariate technique that is concerned with separating distinct set of objects and with allocating new objects into previously defined groups (Johnson & Wichern, 2007) [21]. Discriminant Analysis is a powerful statistical tool that is concerned with the problem of classification. This problem of classification arises when an investigator makes a number of measurements on an individual and wishes to classify the individual into one of the several population groups on the basis of these measurements (Morrison, 1967) [29].

Either Linear or Quadratic rules may be used to derive classification equations in discriminant analysis for the purpose of predicting group membership. Generally, the decision about which rule to use is governed by the degree to which the separate group covariance matrices are unequal (Young, 1993) [40].

Assumptions of Discriminant Function

Once a Discriminant Analysis is contemplated, it may be important to check whether the research data satisfy the assumptions of a Discriminant Analysis. Some of the assumptions to be examined include no outliers, normality assumption and equality of variance covariance matrices.

Multivariate Normality and Outlier Detection

Discriminant Analysis assumes that data for the independent variables represent a sample from a multivariate normal distribution. Many statistical tests and graphical approaches are available to check the multivariate normality assumption. Burdenski (2000) reviewed several statistical and practical approaches, including the Q-Q plot, box-plot, stem and leaf plot, Shapiro-Wilk and Kolmogorov-Smirnov tests to evaluate the univariate normality, contour and perspective plots for assessing bivariate normality, and the chi-square Q-Q plot to check the multivariate normality. In addition, Mardia (1970) [27], Henze and Zirkler (1990) [22], and Royston (1992) proposed various tests for multivariate normality. Holgersson (2006) stated the importance of graphical procedures and presented a simple graphical tool, which is based on the scatter plot of two correlated variables to assess whether the data belong to a multivariate normal distribution or not. For a bivariate data (although could also be used for all $p \geq 2$), a somewhat more formal way for judging the joint normality of a data set is based on the squared generalized distances

$$d_i^2 = \left(X_i - \bar{X}\right)^T \boldsymbol{S}^{-1}\left(X_i - \bar{X}\right) \qquad i = 1, 2, \ldots, n \qquad (1)$$

Specifically, the values of $d_i^2, i = 1, 2, \ldots, n$ are compared with the χ^2 quantiles with p degrees of freedom. In the decision, p-variate normality is indicated if roughly half of the d_i^2 are less than or equal to $q_{c,p}(.50) = 20.3$ where $q_{c,p}\left(\left(i - \frac{1}{2}\right)/n\right)$ is the $100\left(\left(i - \frac{1}{2}\right)/n\right)$ quantile of the chi-square $\left(\chi^2\right)$ distribution. Likewise, with the $d_{(1)}^2 \leq d_{(2)}^2 \leq \ldots \leq d_{(n)}^2$ ordered, the graphical approach also graphs the pairs $\left(q_{c,p}\left(\left(i - \frac{1}{2}\right)/n\right), d_{(i)}^2\right)$, producing a straight line that passes through the origin with slope 1.

Test for Outlier Detection

One of the steps for detecting outliers is to make a scatter plot of each pair of variables. Thus the scatter plots for the various sizes are given in the figures below.

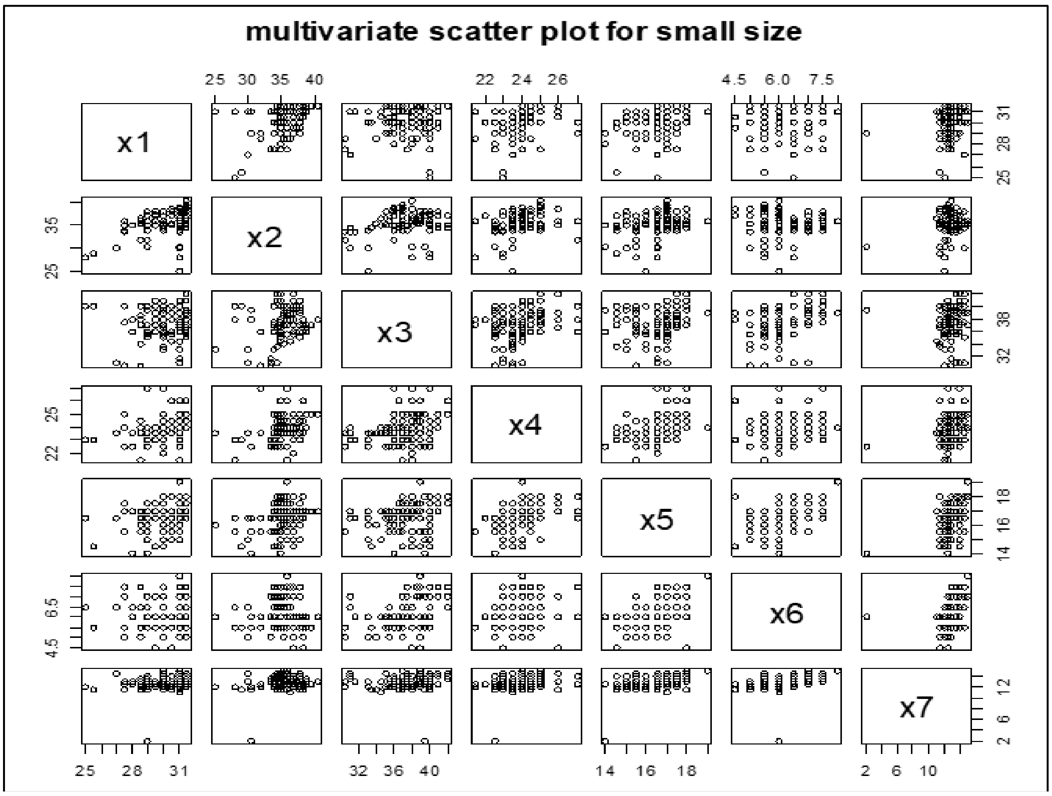

Fig 1: *Two-dimensional multivariate scatter plot for small size*

This plot gives dimensional representation of the variables $\left(X_1, \ldots, X_7\right)$ for the small size. The plot also showed that there are outliers in the data set.

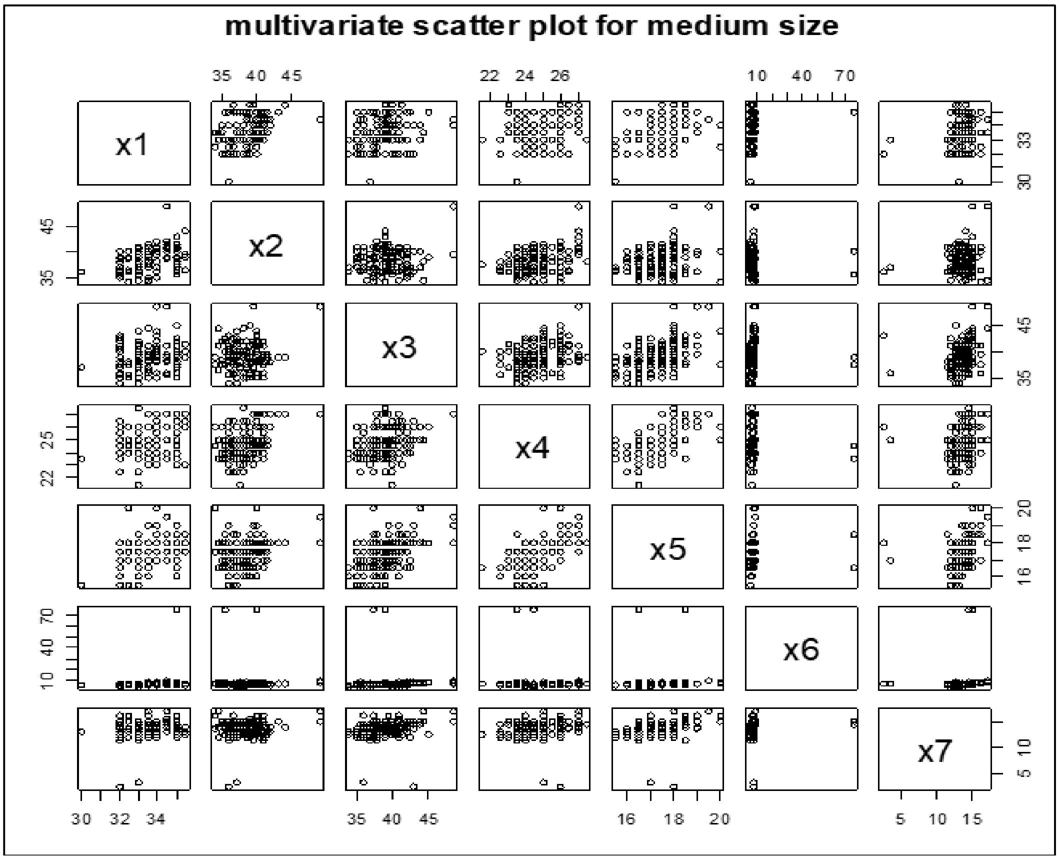

Fig 2: *Two-dimensional multivariate scatter plots for the medium size*

This plot gives dimensional representation of the variables $\left(X_1,...,X_7\right)$ for the medium size. The plot also showed that there are outliers in the data set.

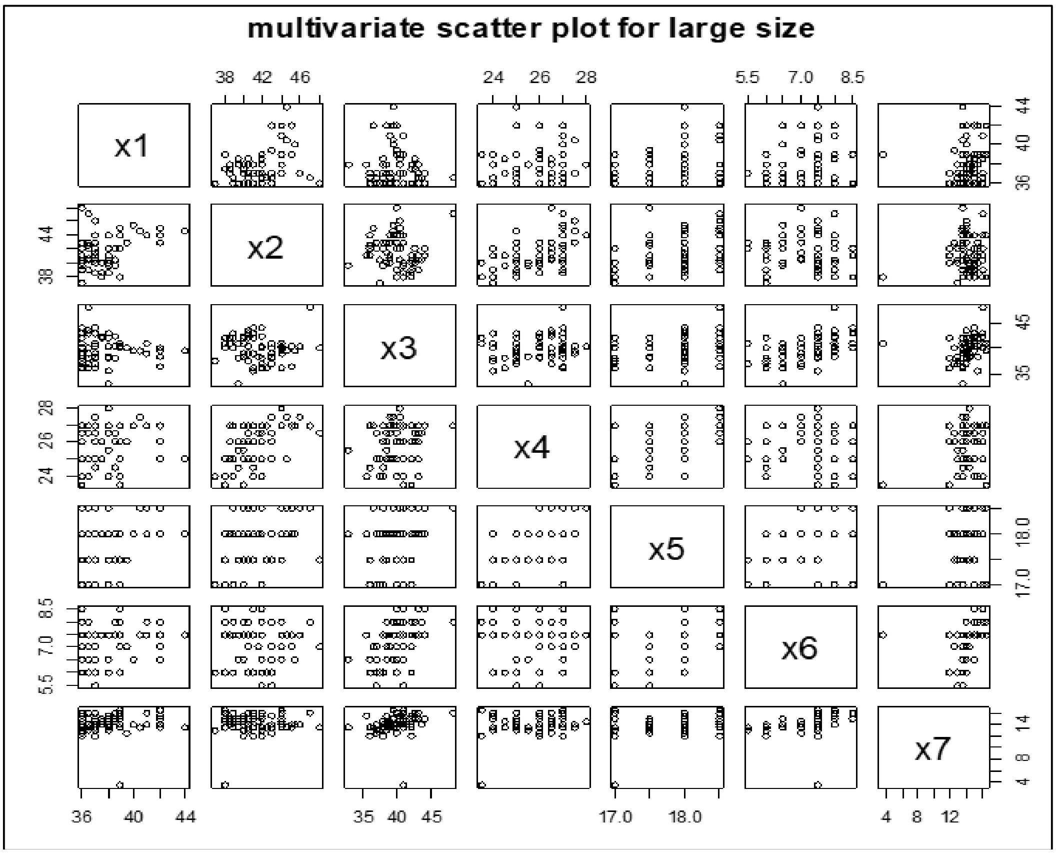

Fig 3: Two-dimensional multivariate scatter plot for the large size

This plot gives dimensional representation of the variables $\left(X_1,\ldots,X_7\right)$ for the medium size. The plot also showed that there are outliers in the data set.

>Squared Distances(nelson.small)
$Squared Distances

```
 [1] 11.724041 11.939374  8.553357 14.514406  8.875050  5.879292 12.570842  7.209104  8.067613
[10]  9.889288  6.644755  3.149976  4.557901  7.975303  4.473226  5.949516  5.682561  1.021645
[19]  5.482490  3.573733 15.208510  9.481026  5.254786  5.596864  3.248441  8.244557  1.629909
[28]  5.347976  6.443889  7.021899  2.052005  8.234515  4.212807  4.459545  4.670194  5.547702
[37]  2.429504  5.749861 14.152824 11.260952 13.129666  4.498031 11.101303  8.374523 12.252344
[46]  8.045354  9.650015  8.543528  3.176002  8.801185  6.431941 10.182072  3.301221  3.542450
[55]  5.671894  2.489490  5.189163  8.940502  7.316416  2.655172  7.657168  5.256383  7.293656
[64]  5.038066  2.381766  3.369887  9.702325  9.592216  5.473515  9.527803  6.380466  4.053240
```

>Squared Distances (nelson. medium)
$Squared Distances

```
  [1]  7.221728 12.574199  5.397572  8.183669  7.265211 16.456234 11.022258  7.971752
  [9]  4.766124  6.011365  4.516669  4.240368 10.630289  8.571689  3.744819  9.837760
 [17]  2.721641  7.267974  3.435358  4.943161  7.401803  4.584455  5.777752  7.146621
 [25] 15.332393 14.398456  7.172346  4.863939  6.294003  6.962396  9.885872  7.796600
 [33]  4.978859  7.861350 10.307776  8.096033  4.583668  6.061001  6.453107 13.063595
 [41] 16.855225  2.761099  4.402071  6.395124  9.734831 11.828865  5.398474  7.178851
 [49]  7.080209  4.205244  5.439610  1.700014  6.235793  7.109094  7.861441  3.005766
 [57]  5.217666  8.093862 15.394242  4.767512  3.225562  5.110959 13.261693  2.927797
 [65]  0.984177  4.230126  4.107091  8.661335  3.009965  1.565201  7.015478  6.482377
 [73]  6.961010  2.850438 10.301747  1.637178  3.529154  2.383950  6.608414  8.271837
 [81]  8.444339  5.652518  4.674068  6.885993  3.585973  4.211033  4.788579  1.269171
 [89]  8.143376  2.125361  1.940444  4.268970 15.311799  5.441972  9.850138 13.894648

 [97]  7.858883  5.805058  8.091549  6.126486  7.074072  6.765511  5.123114  6.608714
[105]  5.250180  6.200494  7.479410  3.971505  7.380388  4.564119 11.660885  3.582159
[113]  8.314681  8.064686  6.312032 14.957777 10.190050  7.082797  3.935239  3.073384
[121]  8.665052  5.662350 15.202392  4.408263  5.227026 10.123572 15.576457  8.736207
[129] 16.675919  7.881750  8.048125  9.123572  7.329869  7.065123  3.716390  4.761946
[137]  8.159410  7.114730  5.694412  3.166300  6.961534  7.157507  4.662171  4.597506
[145]  6.244476  3.028803  8.161490  8.058149  6.720425  9.860008  6.361796  3.043157
[153]  4.236414  2.652671  8.762414  5.163547  7.883383 10.783454  9.279606  5.015656
[161]  6.400413  6.052651 10.609836  6.347508  7.615389  7.294839  8.534155  2.969425
[169]  4.055027  6.783401  6.929100 16.168493  7.111521  3.670760  6.509867  2.927454
[177]  8.403033  7.472458 12.785782  7.831962
```

>Squared Distances (nelson. large)
$Squared Distances

```
 [1] 11.565360 12.745713 11.929763  6.683913  6.046469  7.320516 12.395280  8.575165  3.880496
[10]  2.494879  9.342975 13.037509 14.015833  3.848008  3.123771  2.816886 11.913834  7.065708
[19]  6.080416  2.759302  7.230768  7.048398  7.160482  3.423479  1.988036  6.366039  8.446023
[28]  5.946901  9.673415 14.751116  1.410111  7.409790  7.686317  6.680892  4.656893  5.141355
[37]  3.978651  6.899000  3.610470  3.234394  8.035194 11.695834  3.605442  8.939390  9.562546
[46]  6.604350 12.268372 10.790873  8.925116  1.564187  5.961626  1.068897  5.987001 10.714759
[55]  9.025138 13.344737  7.310839  3.817157  9.811199 14.047888  4.287875  7.344562  5.809443
[64]  1.820933  2.394857  5.057105  5.771171  2.898723  5.471505  5.130965  8.783960  4.409631
[73]  2.778326  8.374574  5.065221  5.871519  9.581510  4.262958  4.097459  1.564187 12.150076
[82]  5.406192 11.724896  3.761182 11.175181  6.771083  8.088336  4.757367  5.760857  4.182701
[91]  5.513756  4.930284 10.942686  5.620653  4.446420  9.690257  8.475958 11.360759
```

Test for Multivariate Normality

Mardia test and chi-square Q-Q plot were used to test for multivariate normality of the sizes (small, medium and large) as follows:

Table 1: Mardia Multivvsriate normality test for the small size group

```
>Mardia Test(nelson.small, qqplot=T)
Mardia's Multivariate Normality Test
---------------------------------------
data             : nelson.small
g1p              : 7.80144
chi.skew         : 93.61728
p.value.skew     : 0.2217146
g2p              : 60.04266
z.kurtosis       : -1.117769
p.value.kurt     : 0.2636658
chi.small.skew : 98.5303
p.value.small    : 0.132857
Result           : Data are multivariate normal.
---------------------------------------
```

Thus, the Mardia Multivariate Normality Test on the small size group data showed a non-significant skewness and kurtosis p-values of 0.2217146 respectively at 0.05 significance levels.

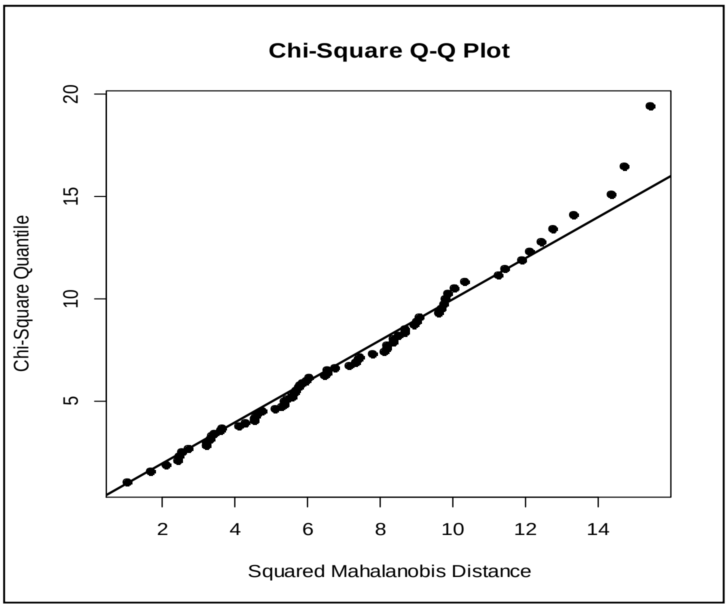

Fig 4: A Chi-square Q-Q Plot showing a multivariate normality for small size group.

Table 2: Mardia Multivariate Normality test for the medium size group

```
>mardia Test(nelson. medium, qqplot=T)
Mardia's Multivariate Normality Test
---------------------------------------
Data              :nelson. Medium
g1p               : 3.506599
chi.skew          : 105.198
p.value.skew      : 0.05871959
g2p               : 60.30721
z.kurtosis        : -1.609251
p.value.kurt      : 0.1075615
chi.small.skew : 107.3963
p.value.small     : 0.04356729
Result            : Data are multivariate normal.
---------------------------------------
```

Thus, the Mardia Multivariate Normality test on the medium size group data showed a non-significant skewness and kurtosis p-values of 0.05871959 and 0.1075615 respectively at 0.05 significance level.

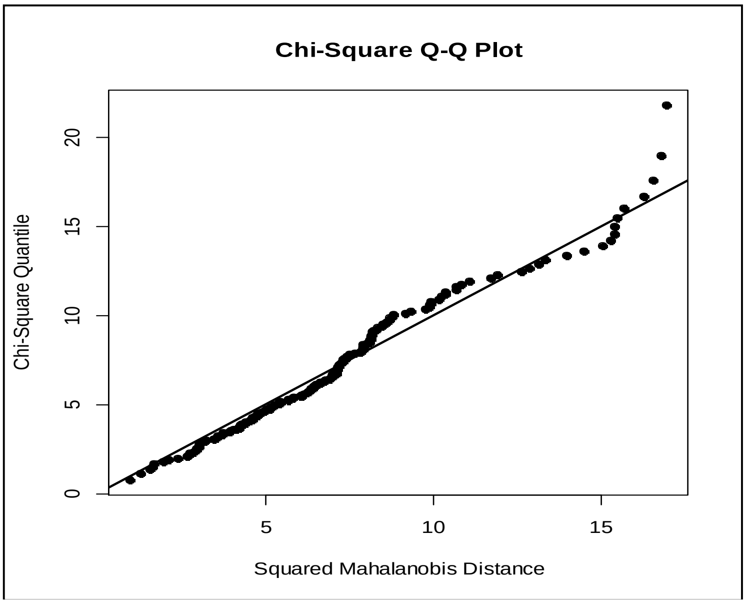

Fig 5: A Chi-square Q-Q Plot showing a multivariate normality for medium size group.

Table 3: Mardia Multivariate Normality test for the large size group

```
>mardia Test(nelson.large, qqplot=T)
Mardia's Multivariate Normality Test
---------------------------------------
Data            :nelson.large
g1p             : 6.180233
chi.skew        : 100.9438
p.value.skew    : 0.1004342
g2p             : 60.57013
z.kurtosis      : -1.071473
p.value.kurt    : 0.283957
chi.small.skew  : 104.8281
p.value.small   : 0.06165799
Result          : Data are multivariate normal.
---------------------------------------
```

Thus, the Mardia Multivariate Normality test on the medium size group data showed a non-significant skewness and kurtosis p-values of 0.1004342 and 0.283957 respectively at 0.05 significance level.

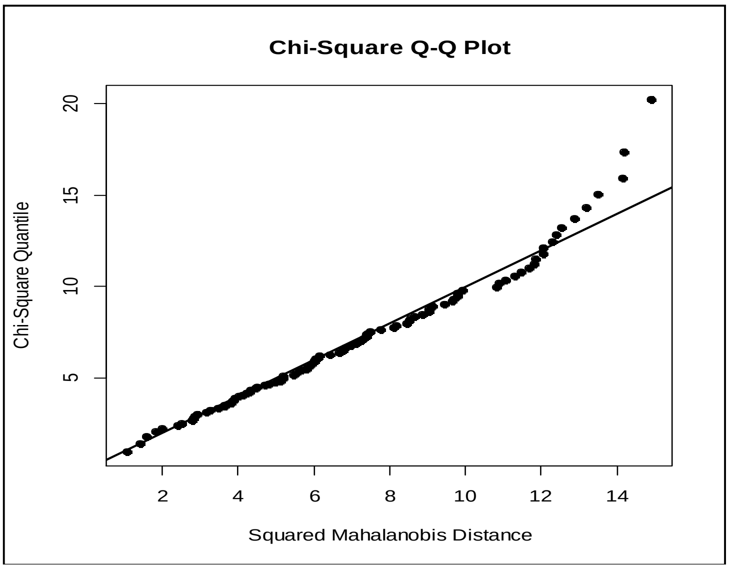

Fig 6: A Chi-square Q-Q Plot showing a multivariate normality for large size group

Equality of Covariance Matrices Using the Box's M Test

Fisher's Linear Discriminant Function assumes that the population covariance matrices across the groups are equal because a pooled estimate of the common covariance matrix is used. To determine whether the covariance matrices for the groups under study are equal, the Box's M test will be used (Box, 1949). The hypotheses: $H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k$ against $H_1 : \Sigma_1 \neq \Sigma_2 \neq \ldots \neq \Sigma_k$ needs to be tested. The test statistic is given by

$$C = (1-u)\left\{\sum_k (n_k - 1)\ln|\boldsymbol{S}_{pooled}| - \sum_k (n_k - 1)\ln|S_k|\right\} \tag{2}$$

Where k is the number of groups, $u = \dfrac{2p^2+3p-1}{6(p+1)(k-1)}\left[\sum_k \dfrac{1}{(n_k-1)} - \dfrac{1}{\sum_k (n_k-1)}\right]$.

The covariance matrices are estimated as follows; the covariance matrix for the small size is given as:

$$s_1 = \begin{pmatrix} 4.0292938 & 2.0110084 & -0.6138009 & 0.10313967 & 0.12299491 & 0.12299491 & 0.34252739 \\ 2.0110084 & 4.6377695 & -1.5831915 & 1.32164026 & 0.33485426 & -0.15856808 & -0.41336072 \\ 2.0110084 & -1.5831915 & 4.1701389 & 0.30178991 & 0.22559664 & 0.84741784 & 1.11169797 \\ 0.1031397 & 1.32164026 & 0.30178991 & 1.22843310 & 0.27626174 & 0.09859155 & -0.02053991 \\ 0.1229949 & 0.33485426 & 0.22559664 & 0.27626174 & 0.16661776 & 0.15375587 & 0.09761346 \\ 0.2582160 & -0.15856808 & 0.84741784 & 0.09859155 & 0.15375587 & 0.56338028 & 0.47652582 \\ 0.3425274 & -0.41336072 & 1.11169797 & -0.02053991 & 0.09761346 & 0.47652582 & 1.08841941 \end{pmatrix}$$

The covariance matrix for the medium size is given as:

$$\boldsymbol{s}_2 = \begin{pmatrix} 1.1807883 & 0.78608007 & 0.72835196 & 0.3118405 & 0.2220670 & 0.2153942 & 0.26474240 \\ 0.78608007 & 3.69909218 & 0.07632682 & 0.9699488 & 0.3959497 & -0.1084264 & 0.09408752 \\ 0.72835196 & 0.07632682 & 5.08233240 & 0.7464385 & 0.8344972 & 1.0127095 & 1.12025140 \\ 0.3118405 & 0.9699488 & 0.7464385 & 1.2794150 & 0.5928771 & 0.2765208 & 0.370561769 \\ 0.2220670 & 0.3959497 & 0.8344972 & 0.5928771 & 0.5768156 & 0.3296089 & 0.39664804 \\ 0.2153942 & -0.1084264 & 1.0127095 & 0.2765208 & 0.3296089 & 0.6716636 & 0.58879578 \\ 0.26474240 & 0.09408752 & 1.12025140 & 0.37056176 & 0.39664804 & 0.58879578 & 1.12045313 \end{pmatrix}$$

The covariance matrix for the large size is given as:

$$\boldsymbol{s}_3 = \begin{pmatrix} 0.86082474 & 0.51472754 & 0.2098675 & 0.2816642 & 0.20029455 & 0.06111929 & 0.15537555 \\ 0.51472754 & 2.20723753 & 0.2587313 & 0.3730276 & 0.07873947 & -0.19666526 & -0.09972649 \\ 0.20986745 & 0.25873133 & 3.7124185 & 0.5216442 & 0.50173575 & 0.68675047 & 0.82871344 \\ 0.28166421 & 0.37302756 & 0.5216442 & 0.6665527 & 0.31080370 & 0.13646644 & 0.22435830 \\ 0.20029455 & 0.07883974 & 0.5017357 & 0.3108037 & 0.77498422 & 0.34751736 & 0.33289501 \\ 0.06111929 & -0.19666526 & 0.6867505 & 0.1364664 & 0.34751736 & 0.47288555 & 0.38020724 \\ 0.15537555 & -0.09972649 & 0.8287134 & 0.2243583 & 0.33289501 & 0.38020724 & 0.76669998 \end{pmatrix}$$

The Pooled Covariance Matrix is given as

$$\boldsymbol{s} = \begin{pmatrix} 1.67418 & 0.96086 & 0.30880 & 0.26070 & 0.19571 & 0.18103 & 0.25009 \\ 0.96086 & 3.47412 & -0.21224 & 0.87505 & 0.29478 & -0.14335 & -0.06392 \\ 0.30880 & -0.21224 & 4.51274 & 0.59262 & 0.61689 & 0.88777 & 1.03710 \\ 0.26070 & 0.87505 & 0.59262 & 1.09766 & 0.44924 & 0.20096 & 0.24967 \\ 0.19571 & 0.29478 & 0.61689 & 0.44924 & 0.54828 & 0.29863 & 0.31764 \\ 0.18103 & -0.14335 & 0.88777 & 0.20096 & 0.29863 & 0.59394 & 0.50752 \\ 0.25009 & -0.06392 & 1.03701 & 0.24967 & 0.31764 & 0.50752 & 1.01501 \end{pmatrix}$$

The determinant of the pooled covariance matrix is: |**S**| = 0.98055

The inverse of the pooled covariance matrix was found with the help of statistical package TORA, and result is given below:

At significance level α, H_0 is not rejected since $C = (241.0105)$ is greater than $\chi^2_{p(p+1)(k-1)/2}(\alpha) = 39.80128$.

$$\boldsymbol{s}^{-1} = \begin{pmatrix} 0.76564 & -0.24235 & -0.00607 & 0.09214 & -0.03930 & -0.20323 & -0.10645 \\ -0.24235 & 0.46488 & 0.03661 & -0.35856 & -0.07162 & 0.26204 & 0.03117 \\ -0.00607 & 0.03661 & 0.33899 & -0.09177 & -0.07436 & -0.30388 & -0.14475 \\ 0.09214 & -0.35856 & -0.09177 & 1.68278 & -1.11992 & 0.05075 & -0.04035 \\ -0.03930 & -0.07162 & -0.07436 & -1.11992 & 3.62927 & -1.17723 & -0.19050 \\ -0.20323 & 0.26204 & -0.30388 & 0.05075 & -1.17723 & 3.86126 & -1.19772 \\ -0.10645 & 0.03117 & -0.14475 & -0.04035 & -0.19050 & -1.19772 & 1.82971 \end{pmatrix}$$

It is important to note that Box's M test is very sensitive to non-normality, such that a significant value indicates either unequal covariance matrices or non-normality or both. Hence, it is necessary to establish multivariate normality before using Box M test (Box, 1949).

Equality of Group Mean Vectors using Wilks Lambda

In discriminant analysis, it is necessary to first test whether or not the mean vectors of the group under study are equal. If the two or three groups differ in their mean vectors, it means that we can construct a discriminant function which hopefully will enable us to distinguish members of one group to those of another group. To test that the mean vector of the three groups (small, medium and large sizes) are equal, the Wilks Lambda statistic (criterion) was used.

Wilks' Test Statistic

The likelihood ratio test of H_0: $\mu_1 = \mu_2 = \ldots = \mu_p$ against H_1: $\mu_1 \neq \mu_2 \neq \ldots \neq \mu_p$ is given as:

$$\Lambda^* = \frac{|E|}{|E+H|}$$

Where, the p×p "hypothesis" matrix H has a between sum of squares on the diagonals for each of the p variables. The p×p "error" matrix E has a within sum of squares for each variable on the diagonal, with analogous sums of products off-diagonal.

We reject H_0 if $\Lambda^* \leq \Lambda^*_{\alpha,p,V_E,V_H}$. Note that rejection is for small values of Λ^*. The parameters in Wilks' Λ^* distribution are:

p = number of variables (dimensions)
k = number of groups (sizes)
V_H = degrees of freedom for hypothesis
V_E = degrees of freedom for error.
The Wilks sampling distribution for multivariate data, for p>1, k=3 are given as:

$$\left(\frac{\sum n_i - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p,2(\Sigma n_i - p - 2)}.$$

Note: For this research work, p=7 (number of variables) and k=3 groups (sizes).

Decision Rule: Reject H_0 if $\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \geq F_{2p,2(\Sigma n_i - p - 2)}$, where $n_i = n_1, n_2, n_3$ (group sizes).

Table 4: MANOVA table for Wilk's test statistic

```
>summary.manova (nelson.manova, test="Wilks")
DfWilksapprox F num DF den DF    Pr (>F)
Sizes        2 0.18404   64.838     14     682 < 2.2e-16 ***
Residuals 347
---
Signif. Codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1
```

The test shows a significant difference of group mean vectors. Thus, the covariance matrices are not equal and Johnson and Wichern (2007) [21] suggested that in such a situation, the use of the quadratic classification rule is advised.

The Quadratic Classification Rule (for $H_1 : \Sigma_1 \neq \Sigma_2 \neq \ldots \neq \Sigma_k$)

The estimate of the quadratic discriminant score $d_i^Q(x)$ is then given as:

$$d_i^Q(\boldsymbol{x}) = -\frac{1}{2}\text{In}\,|\boldsymbol{S}_i| - \frac{1}{2}\left(\boldsymbol{X} - \overline{\boldsymbol{X}}_i\right)\boldsymbol{S}_i^{-1}\left(\boldsymbol{X} - \overline{\boldsymbol{X}}_i\right) + \text{In}p_i$$

And classification rule is based on the sample as follows:

Allocate x to π_i if the quadratic score $d_i^Q(\boldsymbol{x}) = \text{largest of } d_1^Q(\boldsymbol{x}), d_2^Q(\boldsymbol{x}), \ldots, d_p^Q(\boldsymbol{x})$.

Results

Using prior probabilities proportional to the sample sizes [**s = 0.0137, m = 0.0170, l = 0.0500],** the confusion matrix is presented in table 5.

Table 5: Confusion matrix for the quadratic discriminant function

Actual Group	Predicted Group			
	Small size (s)	Medium size (m)	Large size (l)	Total
Small size(s)	72	0	0	72
Medium size (m)	1	174	5	180
Large size (l)	0	3	95	98
TOTAL	73	177	100	350
Correctly classified	72	174	95	341

Thus, from table 5 we obtain the total probability of misclassification, which is the apparent error rate, and is calculated as follows:

$$APER = \frac{(1+5+3)}{350} = 0.025714285$$

Thus, the APER value is about 3% which suggest a relatively low rate of misclassification set.

Partition plots of the various paired anthropometric measurements with prior probability using quadratic discriminant analysis (s = 0.0137, m = 0.0170, l = 0.0500).

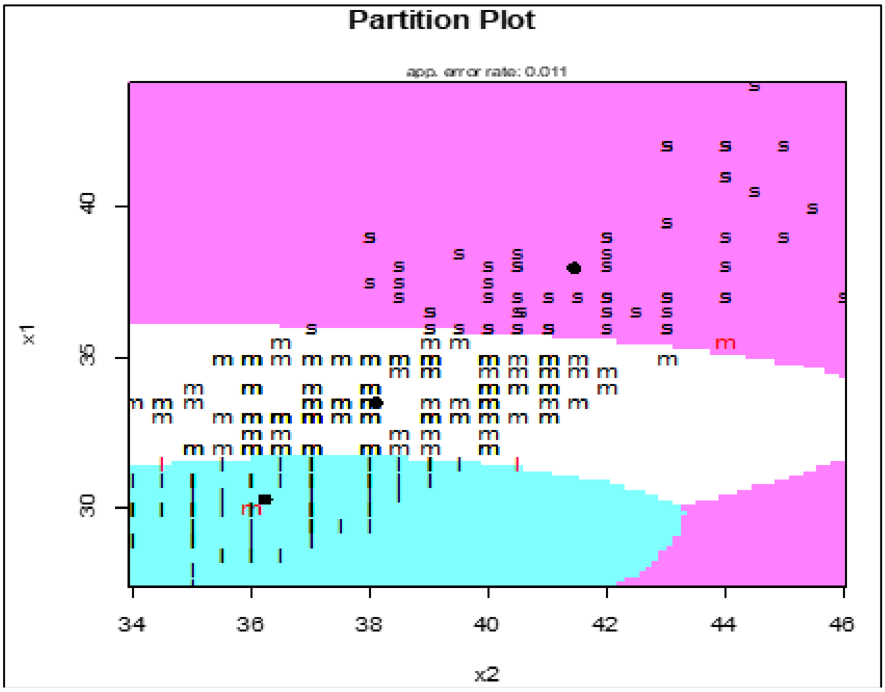

Fig 7: Partition plot of waist (X_1)

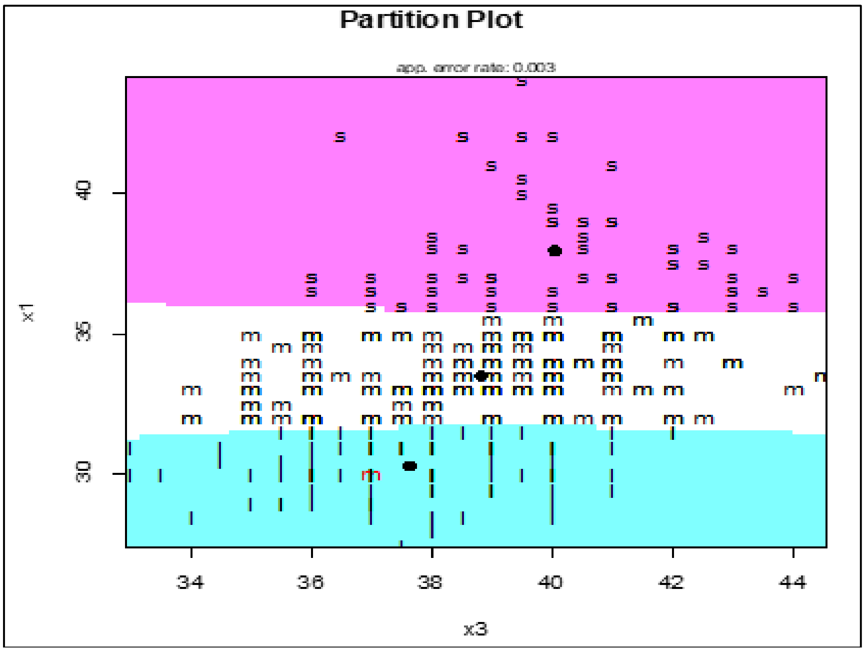

Fig 8: Partition plot of and hip(X_2) waist(X_1) and length (X_3)

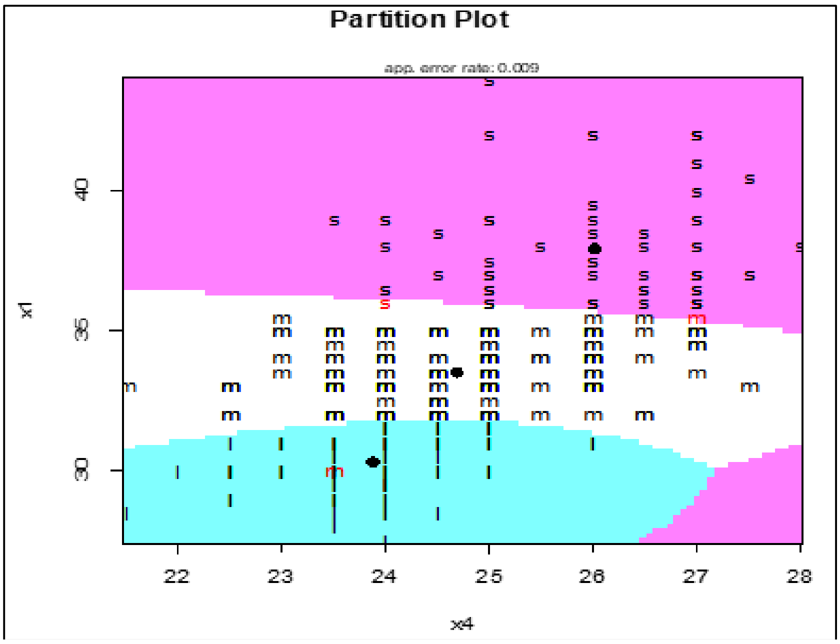

Fig 9: Partition plot of waist(X_1)

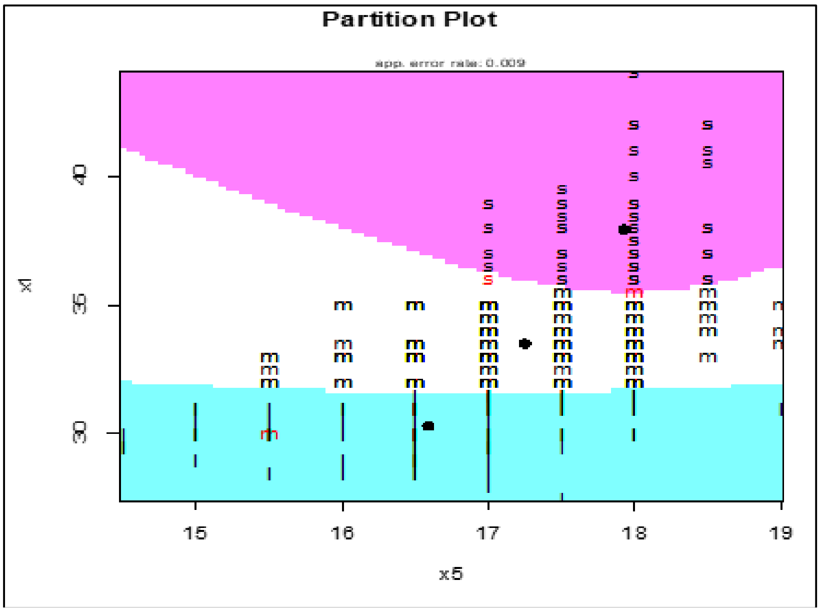

Fig 10: Partition plot of and lap (X_4) waist(X_1) and knee (X_5)

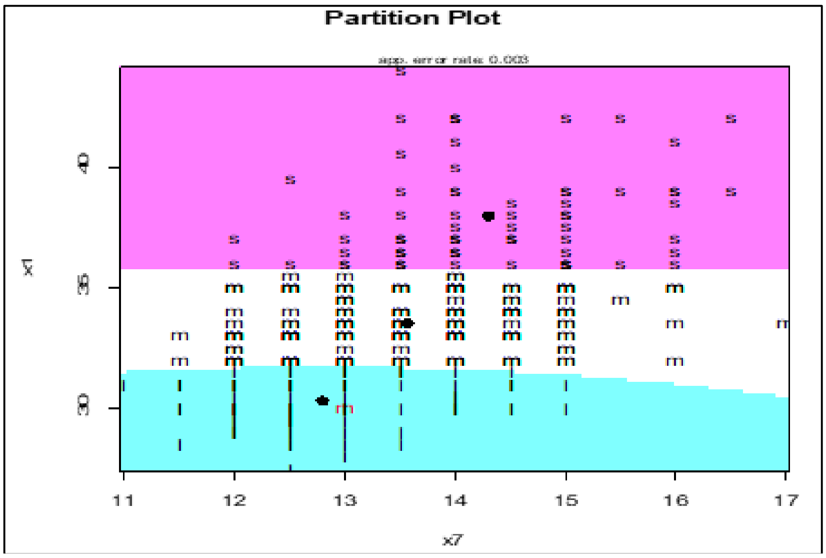

Fig 11: Partition plot of waist (X_1)

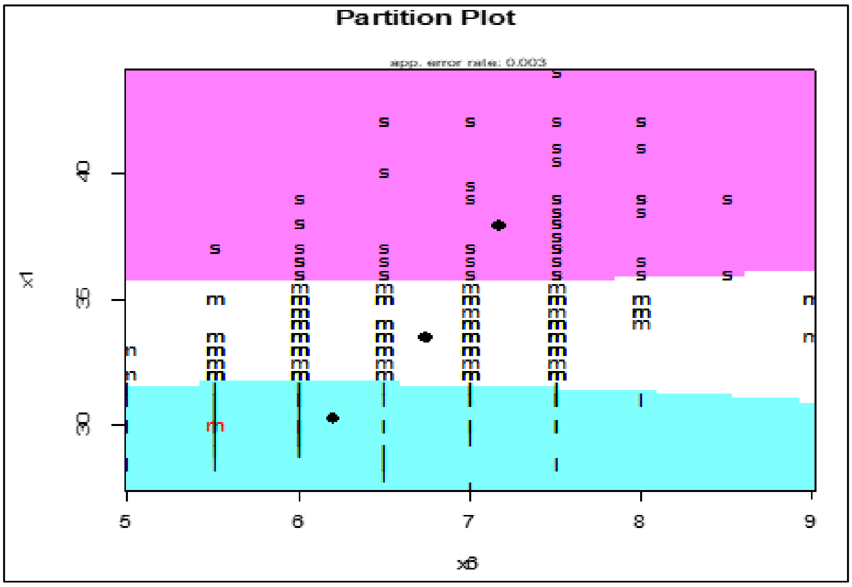

Fig 12: Partition plot of waist and Flap(X_6) waist (X_1)and base (X_7)

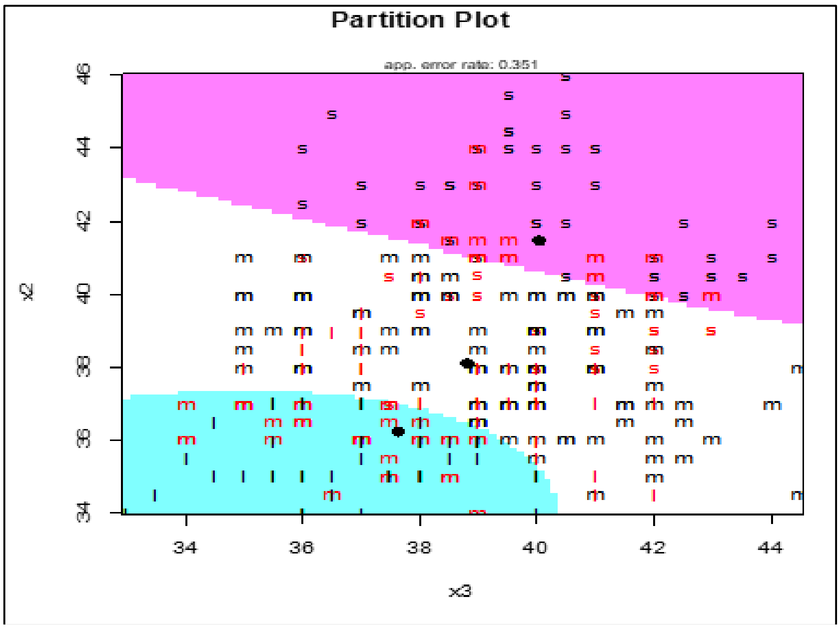

Fig 13: Partition plot of hip (X_2)

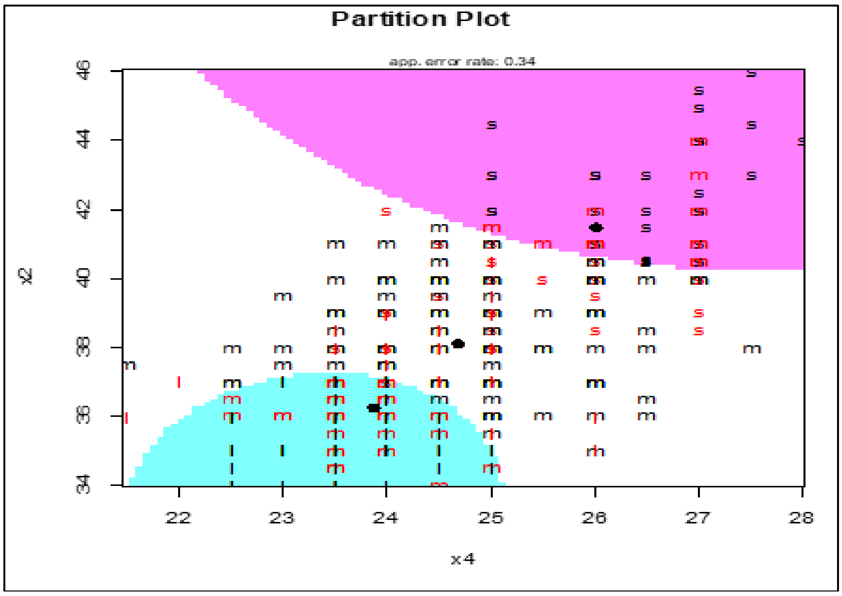

Fig 14: Partition plot of hip And Length (X_3) (X_2) and lap(X_4)

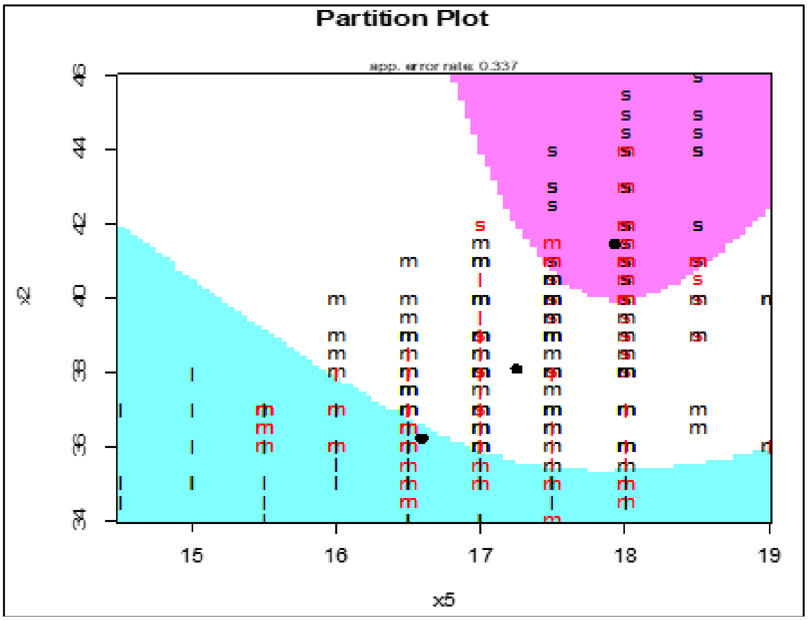

Fig 15: Partition plot of hip (X_2)

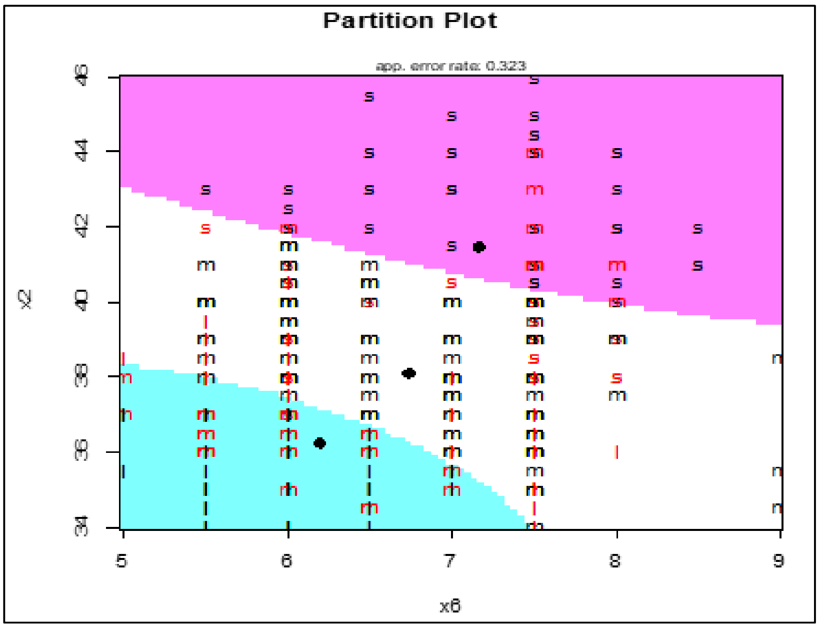

Fig 16: Partition plot of hip and knee (X_5) (X_2) and flap (X_6)

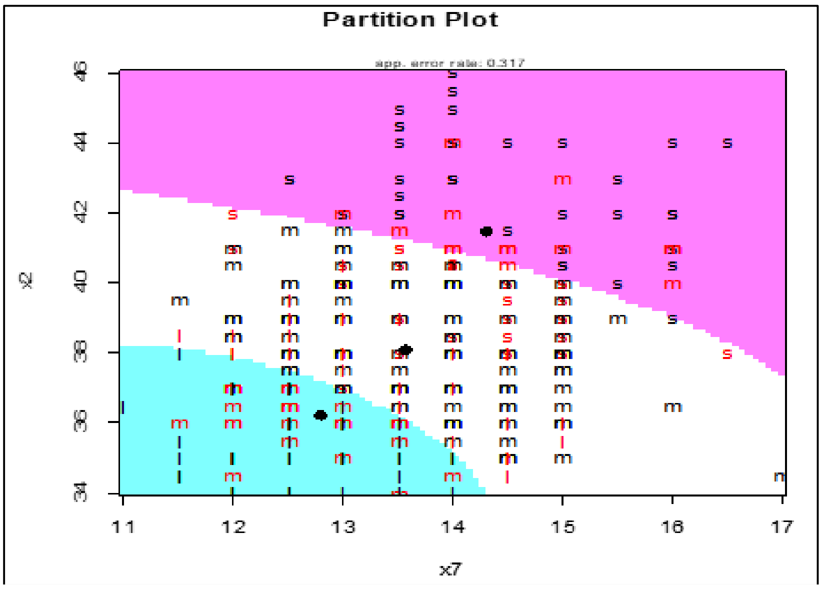

Fig 17: Partition plot of hip

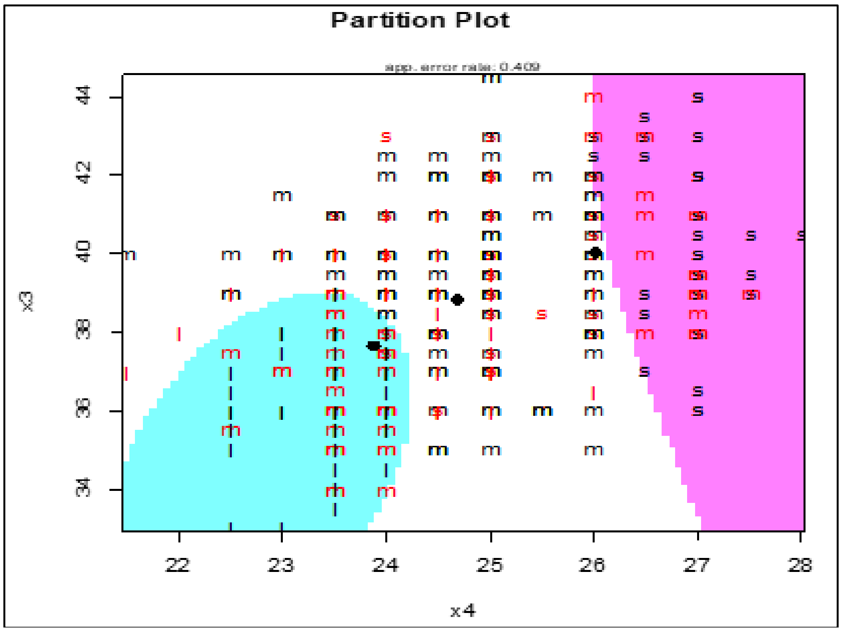

Fig 18: Partition plot of and base (X_7)

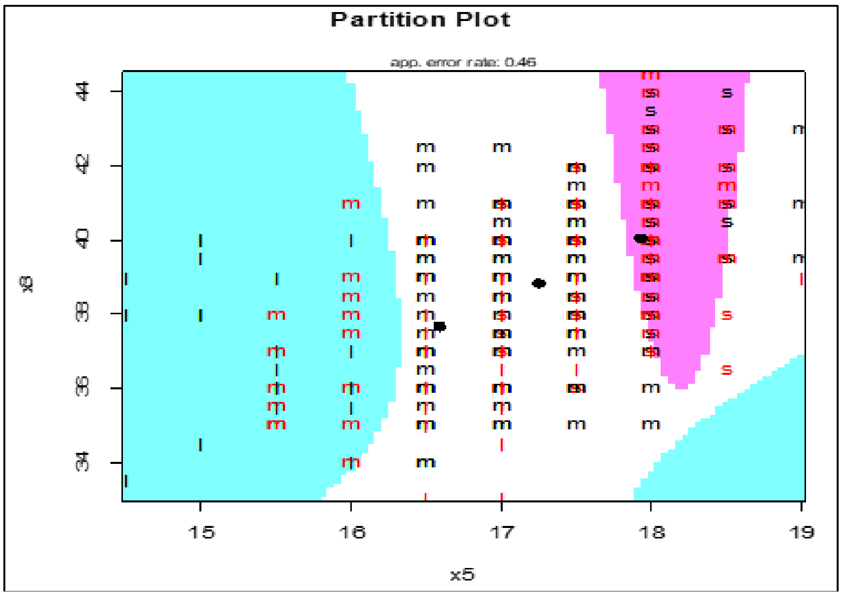

Fig 19: Partition plot of length

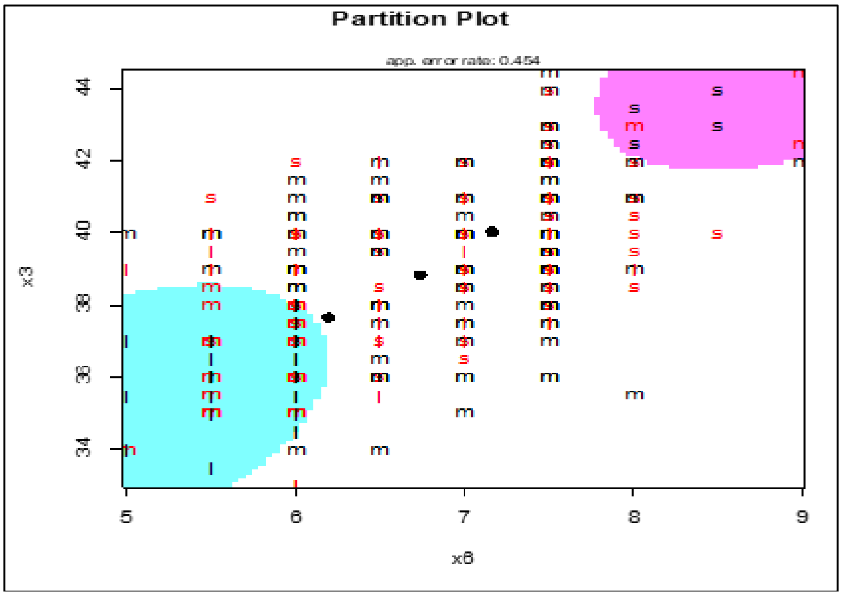

Fig 20: Partition plot of (X_3) and knee (X_5) length (X_3) and flap (X_6)

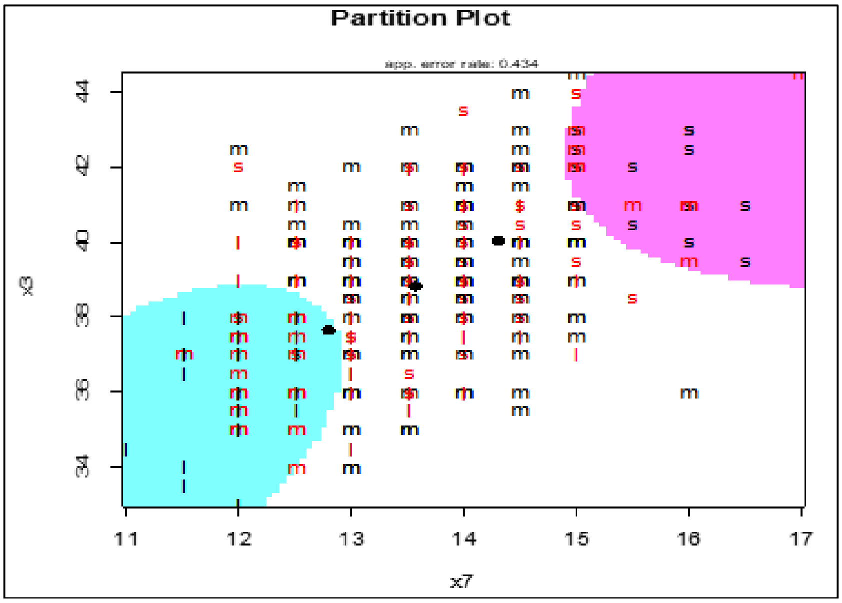

Fig 21: Partition plot of length

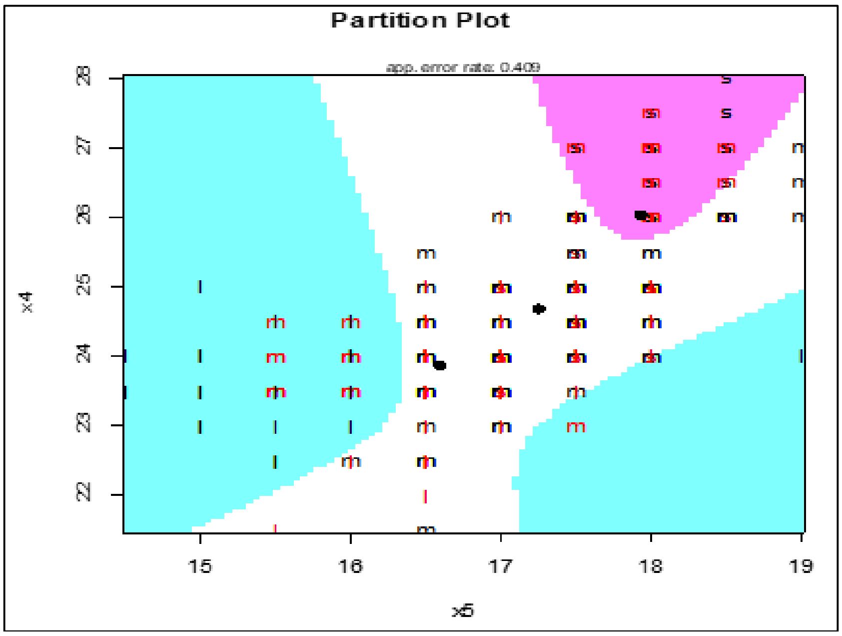

Fig 22: Partition plot of (X_3) and base (X_7) lap (X_4) and knee (X_5)

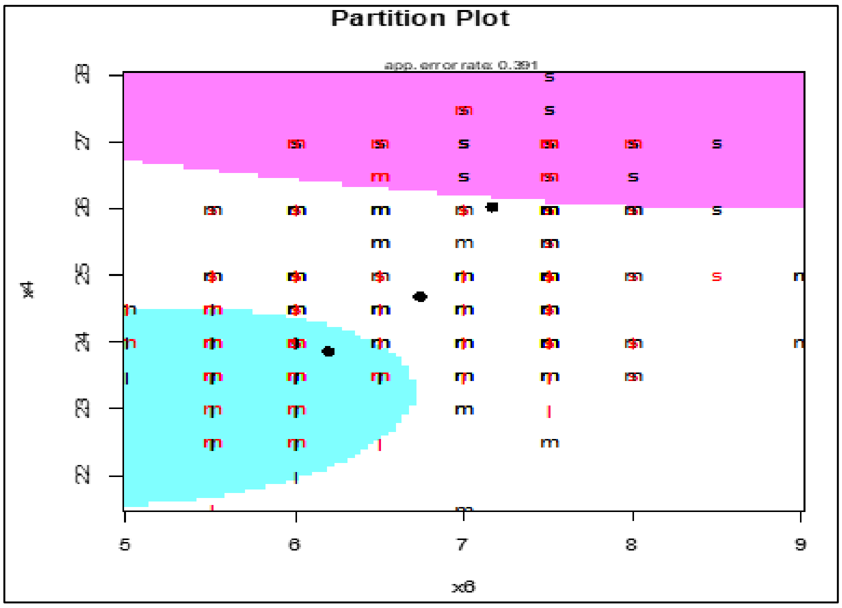

Fig 23: Partition plot lap (X_4)

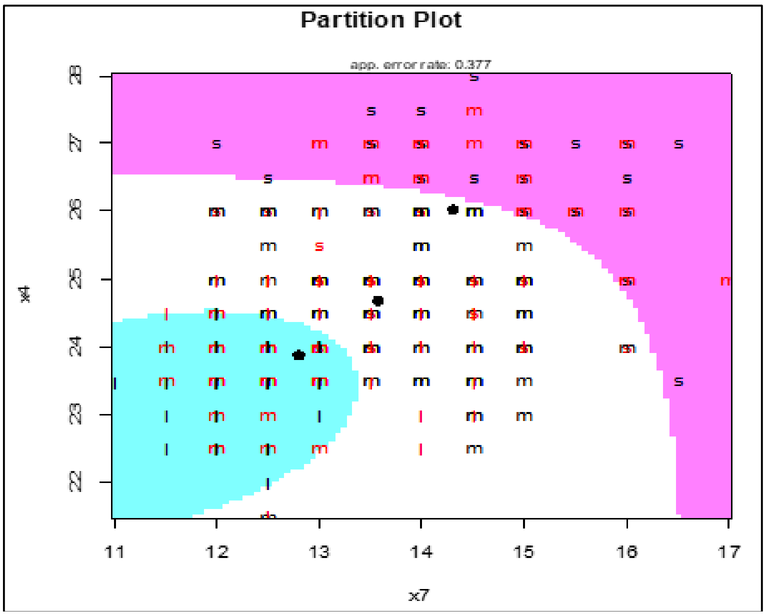

Fig 24: Partition plot of lap (X_4) flap (X_6) base (X_7)

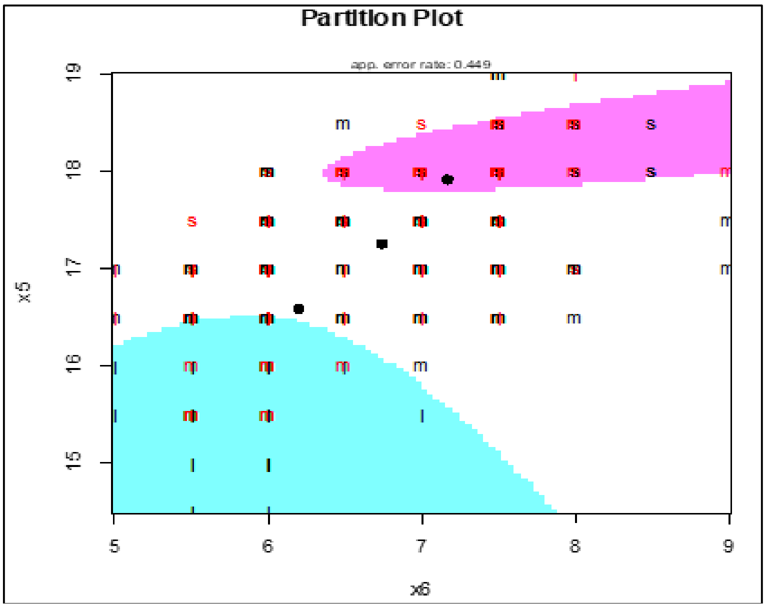

Fig 25: Partition plot knee (X_5)

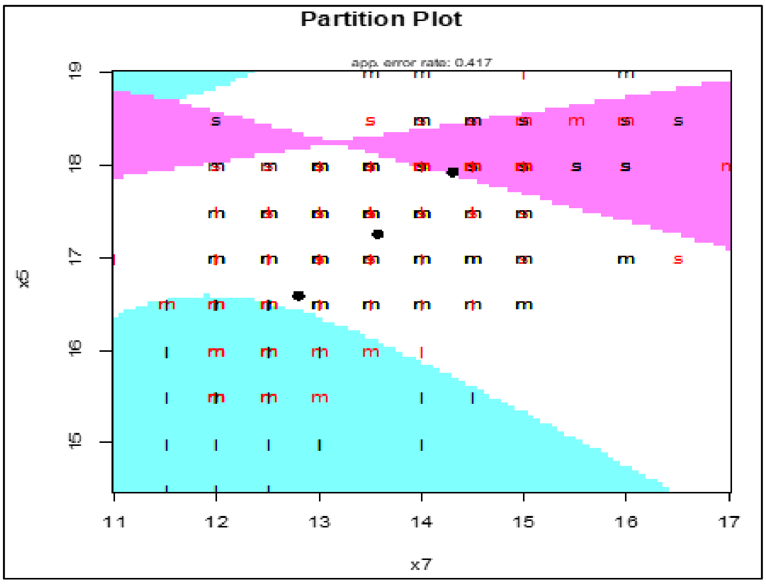

Fig 26: Partition plot of knee(X_5) and flap (X_6) and base (X_7)

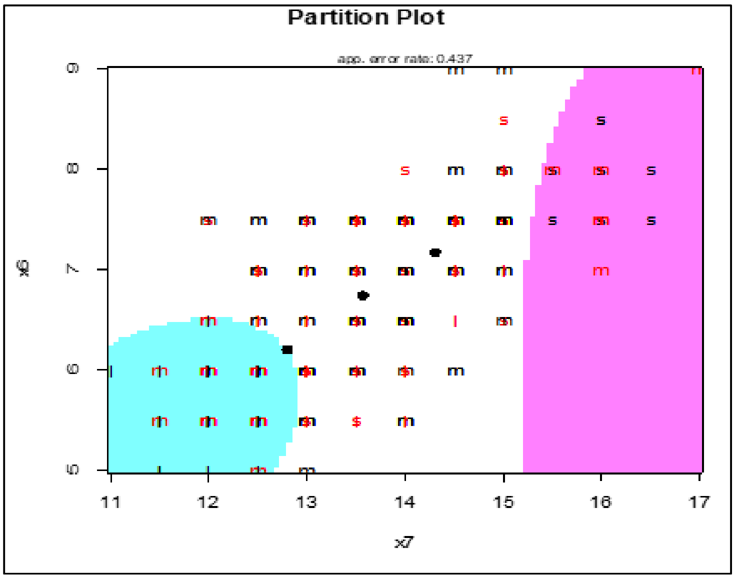

Fig 27: Partition plot of flap (X_6) and base (X_7)

Interpretation of partition plots

The partition plots for quadratic discriminant analysis uses the R Klap package. It maintains the proportions of the data set as the prior probabilities. Thus, the partition plots showed the various paired anthropometric measurements with their prior probabilities (i.e., the probability of misclassifications of one group into another one).

Conclusion

The Quadratic Discriminant Function was applied in this study. The assumptions of normality were established while that of equal covariance matrices was invalidated. Thus, the quadratic classification rule was employed per pair of the variables considered. On the overall, the analysis showed that about 3% of the sizes sampled were misclassified but when the derived classification function was evaluated it correctly classified about 97% of the sizes under study.

References

1. Aldrich W. *History of Sizing and Ready-to-wear Garments. Susan P. Ashdown (ed.), Sizing in clothing Developing Effective Sizing Systems for Ready-to-wear Clothing*, Cambridge: Woodhead Publishing Limited, 2007, 1-56.
2. Anderson TW. *An Introduction to Multivariate Analysis*. John Wiley andSons Inc 1958.
3. Ashdown SP. An Investigation of the Structure of Sizing Systems: A Comparison of the multi-dimension optimized sizing systems generated from anthropometric data with the ASTM standards D5585. International Journal of Clothing Science and Technology 1998;10(5):324-341.
4. Ashdown SP. *Introduction to Sizing and Fit Research*. Research Fit 2000. The Fit Symposium, Clemson Apparel Research, Clemson, SC May 2000.
5. Ashdown SP, Delin MR. Perceptual testing of apparel variation. *Applied Ergonomics* 1995;26(1):47-54.
6. Ashdown SP, Loker S, Adelson C. Use of body scan data to design sizing systems based on target markets 2004. Retrieved: (Jan 12, 2008) from http//:www.cornell.edu/units/txa/research/ntc/SOI-CR01-03.pdf.
7. Audu I, Usman A. An Application to Group Dependent Transformation of Fishers' Discriminant Analysis in a University of Technology. International Journal ofLatest Research in Science and Technology 2013;2(2):40-45.
8. Cattell RB. The Scree Test for the Number of Factors, Multivariate Behavioural Research 1966, *245-276.*
9. Chun-Yoon J, Jasper CR. Consumer Preferences for Size Description Systems of Men's and Women's Apparel. The Journal of Consumer Affairs 1998;29(2):429-441.
10. Cooley WW, Lohnes PR. Multivariate Procedures for the Behavioral Sciences Inc. 1962.
11. Craig HI. *Clothing: A Comprehensive Study*. New York. T.B. Lippincolt 1968.
12. Devarajan P, Istook K. Validation of Female Figure Identification Technique (FFIT) for apparel software, Journal of Textile and Apparel, Technology and Managements 2004;4(1):1-23.
13. Doshi G. Size and Fit Problems with Ready-made Garments 2006. Retrieved (May 20, 2008) from http://www.enzinearticles.com.
14. Faust ME, Carrier S, Baptiste P. Variations in Canadian Women's Ready-to-wear Standard Sizes. Journal of Fashion Marketing and Management 2006;10(1):71-83.
15. Fiore AM, Lee S, Kunz G. Individual Differences, Motivations and Willingness to use Mass Customization Options for Fashion Products. European Journal of Marketing 2004;38(7):835-849.
16. Fisher RA. "The Use of Multiple Measurements In Taxonomic Problems". Annals of Eugenics 1936, 179-188.
17. Frith H, Gleeson K. Clothing and embodiment: men managing image and appearance. Psychology of Men and Masculinity 5(1):40-48.
18. Hogge VE, Baer M, Kang-Park J. Clothing for elderly and non-elderly men: a comparison of preferences, perceived availability and fitting problems, Clothing and Textile Research Journal 1988;6(4):47-53.

19. Honey F, Olds T. The standard Australian Sizing System: Quantifying the Miscach, Journal of Fashion Marketing, 2007;11(3):320-331.
20. ISO 3635 Size Designation of Clothes – Definitions and Body Measurement Procedure; International Organizational Development, IFAS, University of Florida 1977.
21. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis.* Englewood Clifts New Jersey 2007.
22. Johnson KP, Workman E. Effect of Fibre Content Information on Perception of Fabric Characteristics. *Home Economics Journal*, 1990;19(2):132-138.
23. Washington DC. American Home Economics Association. Lachenbrunch P.A. *Discriminant Analysis*.Hafer Press New York 1975.
24. Liu K, Dickerson GG. Taiwanese Male Office Workers: Criteria for business apparel purchase, Journal of Fashion Merchandising and Management 1999;3(3):255-66.
25. Mason AM. *Constraints affecting the growth of the* "Juakali" clothing manufactures in Nairobi-Kenya. M.Sc. Thesis, The Manchester Metropolitan University 1998.
26. Mastamet-mason A. *An explication of the problems with apparel fit experience by female Kenya consumers in terms of their unique body shape characteristics.* Ph.D. Thesis, University of Pretoria 2008.
27. Mardia KV. Measures of multivariate skewness and kurtosis with application. Biometrika, 1970;57(3):519-530.
28. Mardia KV. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies.Sankty a: The Indian Journal of Statistics, Series B (1960-2002), 1974;36(2):115-128.
29. McCormick DP, Kimuyu P, Kinyanjui M. *Weaving through reforms*: Nairobi's small garment producers in a liberalized economy. A Paper presented at the East African Workshop on Business Systems in Africa, Collaborately by the Institute for Development Studies; IDS-University of Nairobi and Centre for Development Research (CDR), Coperhagen, Denmark 2002.
30. Morrisson DF. *Multivariate Statistical Methods*, McGraw-Hill Book Company, New York 1967.
31. Otieno R, Harrow C, Lea-Greenwood G. The unhappy shopper, a retail experience: exploring fashion, fit and affordability. International Journal of Retail and Distribution Management 2005;33(4):298-309.
32. Pisuit G, Connell J. Fit Preferences off Female Consumers in the USA. Journal of Fashion Marketing and Management, 2007;11(3):293-384.
33. Version R. 2.15.2 R Development Core Team, *R* Foundation for Statistical Computing 2012. ISBN 3-9005-07-0, URL: http//:www.R-Project.org.
34. Simmons K, Istook CL, Devaraajan P. Female figure identification technique (FFIT) for apparel. Part 1: describing female shapes. Journal of Textile and Apparel, Technology and Management 2004;4(1):1-16.
35. Sproles GB. *Fashion; Consumer Behaviour Towards Dress*. Minnesota, Buress: Publishing Company 1979.
36. Stamper AA, Sharp S, Donnel LB. Evaluating Apparel Quality. (2nded.) New York Fuirchild Publications 1991.
37. Ulrich PV, Anderson-Connell LJ, Wu W. Consumer Co-design of Apparel for Mass Customization. Journal of Fashion Marketing and Management 2013;8(4):398-412.
38. Workman JE. Body measurement specification for fit models as a factor in clothing size variation. Clothing and Textile Research Journal 1991;10(1):31-36.
39. Yoon JC. A Methodology For Devising an Anthropometric Size Description System for Women's Apparel [CD-ROM]. Abstract from: Proquest file: Discriminant Abstracts Item: ACC9218379 (DAI-13 53/07) 1993.
40. Young B. *Quadratic versus Linear Discriminant Analysis*, A paper presented at the annual meeting of Mid-South Educational Research Association 22nd, New Orleans, LA, 1993;11-12, 1993
41. Zwane PE, Magagular N. Pattern design for women with disproportionate figure; A case study for Swaziland. International Journal of Consumer Studies 2006;31:283-287.