

# Developing Structured Sizing Systems for Manufacturing Ready-Made Garments of Indian Females Using Decision Tree-Based Data Mining

Hina Kausher, Sangita Srivastava

**Abstract**—In India, there is a lack of standard, systematic sizing approach for producing readymade garments. Garments manufacturing companies use their own created size tables by modifying international sizing charts of ready-made garments. The purpose of this study is to tabulate the anthropometric data which cover the variety of figure proportions in both height and girth. 3,000 data have been collected by an anthropometric survey undertaken over females between the ages of 16 to 80 years from the some states of India to produce the sizing system suitable for clothing manufacture and retailing. The data are used for the statistical analysis of body measurements, the formulation of sizing systems and body measurements tables. Factor analysis technique is used to filter the control body dimensions from the large number of variables. Decision tree-based data mining is used to cluster the data. The standard and structured sizing system can facilitate pattern grading and garment production. Moreover, it can exceed buying ratios and upgrade size allocations to retail segments.

**Keywords**—Anthropometric data, data mining, decision tree, garments manufacturing, ready-made garments, sizing systems.

## I. INTRODUCTION

In earlier times, both men and women used to wear draped garments like dhoti, sari, gamucha, lungi, etc... Gradually, some garments had been started to stitch. Women used to stitch garments like blouse, petticoat at home. Afterwards some people called as “tailors” started stitching of customized garments. They used to complete the whole pattern making, cutting and construction independently or with assistants in their shop.

When the industrial revolution started and the mass produced clothing was introduced by the apparel industries in India, the first and biggest challenge before the apparel industries was the ‘sizing of the mass produced garments’; to solve these issues they started using international sizing charts for making readymade garments for Indian people.

As time went by, apparel manufacturers made some modifications in the size charts of other countries and used these charts for making structured garments for Indians. When manufacturers make garments using these international charts, they fail to consider the reality that Indian people are far different in heights, weights, proportions, and body types from the U.S and U.K people. Unfortunately, these modified sizing

charts could not justify the Indian anthropometry, resulted in dissatisfaction of customers with the fit of ready to wear clothing. The problem still remains the same as it was before. Still we Indians do not have our own original size table for making ready to wear garments. The average height of women of United States is much higher than the Indian women. According to [3], the national health statistics reports 2018, of the United States; the average height of U.S. women is currently 5 feet 4 inches while the mean height of Indian women lies between 5 inches to 5 feet one inch. These differences exclude the majority of women from buying their clothes. As [1] states, a nightmarish situation for the consumer would be created. Consumers can no longer rely on the size label to identify a fitting garment and must therefore spend an undue amount of time finding the right one. Reference [4] shows that India is poised on the brink of a major retail revolution. Dressing habits are being changed by the younger generation and western clothes are becoming more and more popular in the cities as well as suburbs. It is at this stage that an efficient garment sizing system is highly needed.

The purpose of this study is to conduct an anthropometric survey, analyse the collected data and develop a sizing system for pattern making of women’s kurta.

## II. LITERATURE REVIEW

Reference [8] defined the sizing system as a table of numbers which presented the value of each of the body dimensions used to classify the bodies available in the population for each size group in the system. But if the number of size groups is large in sizing system, each group will have a small number of individuals who will be nearly the same to one another in body measurements; therefore, all individuals in that size group will get a very good fit. But large number of size groups may actually cause customers to be dissatisfied with the shopping experience because of confusion about size selection and the necessity to try on too many garments in order to get the right size [3].

Reference [7] defined the aim of an apparel sizing system as to divide a heterogeneous population into homogeneous subgroups. Members having similar size and shape are to be put in one subgroup so that the garments of respected group can adequately fit each of them. Different subgroup members are dissimilar, therefore, would require different sizes of garments

Reference [5] mentioned that India does not have its own nationalised standard sizing system for making ready-made

Hina Kausher, Research Scholar, is with the Department of Home Science, University of Allahabad, Prayagraj, India (phone: 6393540831, e-mail: elf.hina@gmail.com).

Prof. Sangita Srivastava, Head, is with the Department of Home Science, University of Allahabad, Prayagraj, India.

garments. There are many reasons behind it: One is that the traditional Indian garments comprised primarily of draped wear and some stitched garments local tailors were available; second, the lack of organized retail sector; third, Indian population was not fashion and fit conscious. In the course of time, people clothing preferences have been shifted from Indian to western wear. Therefore, the need of sizing system is becoming more and more acute.

Reference [5] also stated that there are as many sizes and sizing systems exist in India as there are brands. Each brand has different sizes, different nomenclature and even different size intervals. Consumers have to try innumerable clothes in each brand for getting correct one.

### III. MATERIALS AND METHODS

This study was divided into following three parts:

#### A. Anthropometric Survey

1,000 women between the ages of 16 and 80 years old were selected for the study. Stratified random sampling was used for collecting the data. Measurements of round bust, round waist, round stomach, round hip, height and weight were taken as per the ISO 8559/1989 body measurements standard. Stadiometer and digital weight machine were used for taking height and weight of the subject respectively.

#### B. Anthropometric Data Analysis

After collecting the data through anthropometric survey, it was analysed by descriptive statistics (mean, median and standard deviation). Subsequently, factor analysis and decision tree technique were implemented for developing the sizing system. IBM SPSS Statistics Version 20, statistical software was used for conducting all the analysis. The purpose of employing each technique is as follows:

##### 1. Principal Component Analysis (PCA) Method

Control body dimensions were selected by reducing the other variables using PCA technique. Selected key dimensions were used to classify the sample population. PCA and Varimax rotation were used for extracting the key body

dimensions.

##### 2. Decision Tree Method

Decision tree algorithm was used for classifying the data through selected key dimensions by PCA technique. In this study, weight was taken as dependent variable whereas bust girth was taken as independent variable.

##### C. Sizing System Development

The purpose of a sizing system is to divide the whole population into homogeneous subgroups in a way that members having similar size and shape are to be put in one subgroup so that the garments of respected group can adequately fit each of them. Since, members of a group are different from the members of other group in terms of body sizes; therefore, garments sizes for each group are different from other groups. In this study, sizing system which was based on size-interval had been developed with the help of obtained clusters through decision tree method.

### IV. RESULTS AND DISCUSSION

#### A. Factor Analysis

Factor analysis was done as [9]. Latent root, percentage of variance and scree test criterion were used to recognize the number of components to be kept and based on the relevant components and factor loadings control dimensions were selected. Therefore, three steps were used for selecting the factor loadings of each component. These were:

##### Step1: Retaining the Factor Components

##### Test of Multicollinearity and Pattern of Relationship

Correlation matrix was used to check the pattern of relationships between the variables. Since, for these data, the value of determinant of the correlation matrix is 0.004 ( $> 0.00001$ ), therefore; multicollinearity is not a problem for these data. All variables correlate fairly well and none of the correlation coefficients are particularly large; therefore, elimination of any variable is not needed at this stage.

TABLE I  
KMO AND PCA RESULTS FOR WOMEN AGED 16 TO 80 YEARS

Age	Kaiser Meyer Olkin (KMO)	Bartlett's Test	Component Factor Extracted (PCA)	Factor	Eigen Values	Scree Plot Criterion	Cumulative Total Variance %
16-80	0.736	5591.789	2	Factor 1 Factor 2	3.680 1.457	2 components	73.39

Table I shows the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity. KMO and Bartlett test statistics values are used to check the suitability of data for doing factor analysis. In this study, the KMO value is 0.736 ( $> 0.5$ ), which falls into the range of being good, therefore factor analysis is appropriate for these data.

Bartlett's measure tests whether the original correlation matrix is an identity matrix or not. For factor analysis, some relationships between the variables are needed and if the R-matrix was an identity matrix then all correlation coefficients would be zero. Therefore, this test is significant if its significance value is less than 0.05 [2]. In this study, Bartlett's

test is highly significant ( $p < 0.001$ ), which shows that factor analysis is appropriate or all the dimensions are suitable for factor analysis.

##### Step2: Factor Extraction

The results of the extracted components from PCA technique show that two components are extracted among the seven components. The eigenvalues associated with each factor represent the variance explained by that particular linear component [2]. Here, factor 1 explains 52.58% of total variance and factor 2 explains 20.8% of total variance. Hence, from these findings, 73.39% variance is explained by two

principal components.

TABLE II  
EXTRACTED COMPONENTS FROM PCA TECHNIQUE

Component	Initial Eigen values			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	3.680	52.576	52.576						
2	1.457	20.811	73.387						
3	.861	12.305	85.692						
4	.662	9.461	95.153	3.680	52.576	52.576	2.881	41.151	41.151
5	.174	2.481	97.634	1.457	20.811	73.387	2.257	32.236	73.387
6	.094	1.336	98.970						
7	.072	1.030	100.000						

Table I shows that two components with eigenvalues  $> 1$  are retained. Also using scree plot criterion (Fig. 2), two components are retained: component 1 and 2. The third criterion to determine the retaining values is percentage of accumulated variance on eigenvalue. In this study, component 1 (PC1) shows eigenvalue of 3.680 and component 2 (PC2) shows eigenvalue of 1.457. Therefore, the cumulative % variance from initial eigenvalue is 73.387.

After determining the retaining components, all factor loadings of each variable in each component are examined.

### Step3: Result of PCA

Anthropometric variables having factor loadings more than 0.4 are clustered within component 1 (PC1) and component 2 (PC2). Seven variables are loaded on two components. In PC1, weight, bust girth, waist girth, height and age whereas in PC2, waist girth, stomach girth and hip girth found with factor loadings more than 0.4. Five variables are found to have high factor loadings ( $\geq 0.70$ ), which are weight, bust girth, waist girth, stomach girth and hip girth.

Round bust, round waist and round hip are most frequently used in patternmaking of kurta for women.

Variables	Factor 1	Factor 2
Weight	0.862*	
Bust girth	0.847*	
Waist girth	0.825*	0.422
Height	0.624	
Age	0.523	
Stomach girth		0.926*
Hip girth		0.912*

\* Marked loadings are  $> 0.7$

The result of the factor analysis shows that besides weight, two anthropometric variables most closely correlate with the PC1 are bust girth and waist girth. Since, bust girth is the anthropometric variable most closely correlating with the PC1 and it is also the most important variable in establishing sizing system for making upper body garments of women. Therefore, bust girth is selected to represent the component 1.

Although stomach girth is the anthropometric variable most closely correlated with PC2, hip girth is more important in developing sizing systems for making women's kurta. Therefore, hip girth is selected to represent the component 2.

Hence, round bust and round girth are selected by factor analysis to be the most important sizing variables.

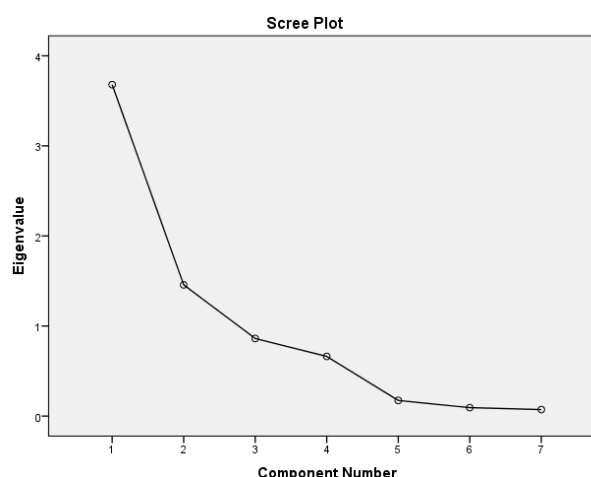


Fig. 1 Scree plot of the eigenvalues of factors

### B. Decision Tree Data Mining

TABLE IV  
CLUSTERS OF UPPER BODY OF 16-80 YEARS OLD WOMEN

Clusters	classification	Frequency
1	Bust girth $\leq 30$ inches	109
2	30.1 < Bust girth $\leq 31$ inches	93
3	31.1 < Bust girth $\leq 32.5$ inches	199
4	32.6 < Bust girth $\leq 33.5$ inches	113
5	33.6 < Bust girth $\leq 34.5$ inches	79
6	34.6 < Bust girth $\leq 35.5$ inches	107
7	35.6 < Bust girth $\leq 37$ inches	101
8	37.1 < Bust girth $\leq 39$ inches	103
9	39.1 < Bust girth inches	94

In this study, CART decision tree technique is used for data mining. IBM SPSS Statistics Version 20, statistical software is used for analysing the data. Weight is taken as the dependent variable whereas; round bust and round hip are taken as independent variable for classifying the dependent variable.

In parent node, the minimum number of samples is 100 and in child node the minimum number of samples is 50.

Fig. 2 shows the results of decision tree analysis. A very first node or parent node is split according to bust girth, resulting in the first level. There are total 9 nodes in first level. The first level is split according to bust girth. In second level, hip girth is used for splitting. At last, nodes which are

generated at the first level are selected for making the sizing system.

Table IV identifies type of women’s figures. Figure type 1

has the smallest bust size ( $\leq 30$  inches) with sample size of 109 whereas figure type 9 has the largest bust girth ( $> 39$  inches) with the sample size of 94.

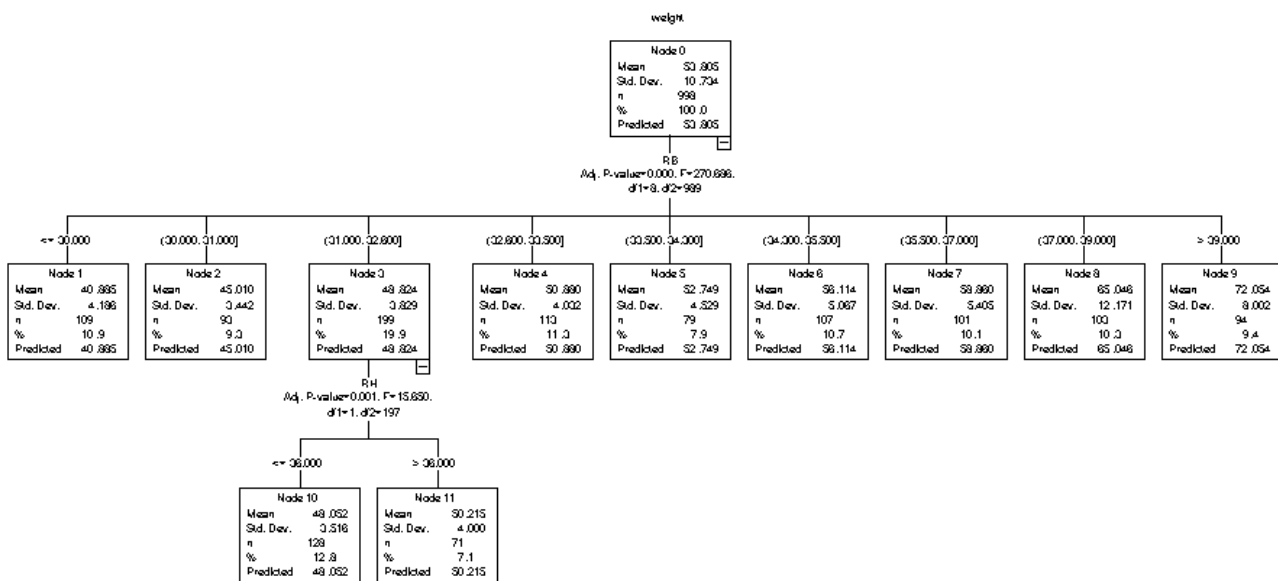


Fig. 2 Decision tree branching

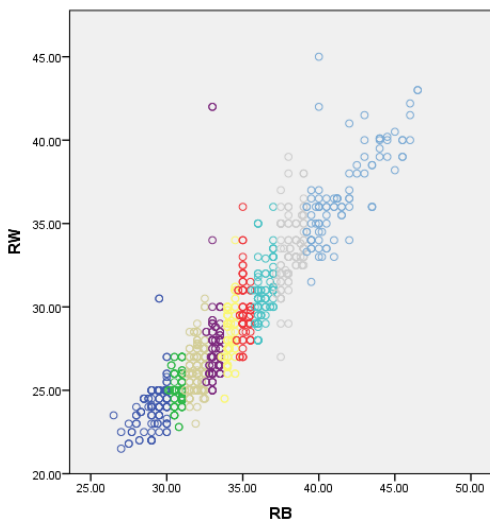


Fig. 3 Scattered plot between round bust and round waist for nine figure types

Fig. 3 represents a scattered plot of bust girth against the waist girth. X-axis presents the bust girth whereas y axis presents the waist girth. Waist girth is also an important body dimension for making upper garments and it significantly correlates with the girth factor.

Fig. 4 represents a scattered plot of bust girth against the hip girth. X-axis presents the bust girth whereas y axis presents the hip girth.

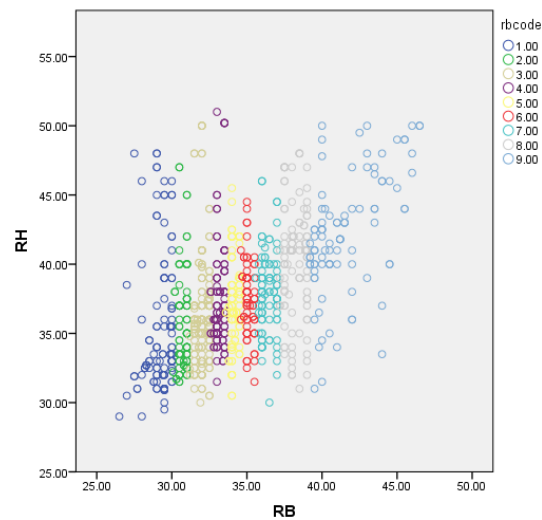


Fig. 4 Scattered plot between round bust and round hip for nine figure types

C. Development of Sizing System

Sizing system was developed using the methods [6]. The most important variables used for making women’s upper body garments were bust, waist and hip girth. Fig. 5 represents distribution plots of round bust on x-axis and round waist on y-axis at the intervals of 2 inches.

Fig. 5 shows a scatter plot for figure type 1, in which x-axis presents the bust girth and y-axis presents the waist girth. The bust girth for type 1 figure is 27, 29 and 30 inches. The other figure types represent the similar scatter plot as figure type 1.

Table IV shows a sizing system for all 9 type of women’s

figure with covering 100% of the sample population.

Total	1	7	18	52	31	109
30			2			2
29						0
28					2	0
27						2
26			15		2	2
25					8	23
24	1	7	8	18	13	39
23			10	17	6	41
	26	27	28	29	30	Total

Bust girth (in inches)

Fig. 5 Scatter plot for cluster 1 (figure type 1) size

A simple sizing system can be made using Table V. Hip girth: 32 < hip girth < 50 was excluded. 40 women out of 998 (4%) have hip girth less than 30 inches whereas, 5 females out of 998 (0.5%) have hip girth larger than 50 inches. Thus, total 4.5% females were excluded from sample population for developing sizing system.

V. CONCLUSION

This paper presented a sizing system with key body dimensions for making pattern of women's kurta. Multivariate analysis using PCA shows that bust girth, waist girth and hip girth are most influential in size determination.

TABLE V  
THE SIZING SYSTEM FOR WOMEN'S KURTA WITH 9 FIGURE TYPES

Figure type	%	N	Waist girth (in inches)	Bust girth (in inches)	Hip girth (in inches)											
					30	32	34	36	38	40	42	44	46	48	50	51
1	10.9	109	24	≤ 30	30	32	34	36	38	40	42	44	46	48	50	51
2	9.3	93	25.5	30.1-31		32	34	36	38	40	42	44	46			
3	19.4	194	26.5	31.1-32.5	30	32	34	36	38	40	42	44	46	48	50	
4	11.7	117	28	32.6-33.5		32	34	36	38	40	42	44	46	48	50	51
5	11.0	110	29	33.6-34.5		32	34	36	38	40	42	44	46			
6	7.6	76	30	34.6-35.5			34	36	38	40	42	44				
7	10.1	101	31	35.6-37			34	36	38	40	42	44	46			
8	10.3	103	34	37-39			34	36	38	40	42	44	46	48		
9	9.4	94	37	> 39				36	38	40	42	44	46	48	50	

TABLE VI  
SIMPLIFIED SIZING SYSTEM FOR PATTERNMAKING OF WOMEN'S KURTA

Figure type	Waist girth (in inches)	Bust girth (in inches)	Hip girth (in inches)										
			32	34	36	38	40	42	44	46	48	50	
1	24	≤ 30	≤ 30 XXXS	≤ 30 XXS	≤ 30 XS	≤ 30 S	≤ 30 M	≤ 30 L	≤ 30 XL	≤ 30 XXL	≤ 30 XXXL	≤ 30 XXXXL	
2	25.5	30.1-31	30.1-31 XXXS	30.1-31 XXS	30.1-31 XS	30.1-31 S	30.1-31 M	30.1-31 L	30.1-31 XL	30.1-31 XXL			
3	26.5	31.1-32.5	31.1-32.5 XXXS	31.1-32.5 XXS	31.1-32.5 XS	31.1-32.5 S	31.1-32.5 M	31.1-32.5 L	31.1-32.5 XL	31.1-32.5 XXL	31.1-32.5 XXXL	31.1-32.5 XXXXL	
4	28	32.6-33.5	32.6-33.5 XXXS	32.6-33.5 XXS	32.6-33.5 XS	32.6-33.5 S	32.6-33.5 M	32.6-33.5 L	32.6-33.5 XL	32.6-33.5 XXL	32.6-33.5 XXXL	32.6-33.5 XXXXL	
5	29	33.6-34.5	33.6-34.5 XXXS	33.6-34.5 XXS	33.6-34.5 XS	33.6-34.5 S	33.6-34.5 M	33.6-34.5 L	33.6-34.5 XL	33.6-34.5 XXL			
6	30	34.6-35.5		34.6-35.5 XXS	34.6-35.5 XS	34.6-35.5 S	34.6-35.5 M	34.6-35.5 L	34.6-35.5 XL				
7	31	35.6-37		35.6-37 XXS	35.6-37 XS	35.6-37 S	35.6-37 M	35.6-37 L	35.6-37 XL	35.6-37 XXL			
8	34	37-39		37-39 XXS	37-39 XS	37-39 S	37-39 M	37-39 L	37-39 XL	37-39 XXL	37-39 XXXL		
9	37	>39			> 39 XS	> 39 S	> 39 M	> 39 L	> 39 XL	> 39 XXL	> 39 XXXL	> 39 XXXXL	

Different figure types are identified using decision tree data mining technique. A sizing system is developed for making kurta for women for the age of 16 to 80 years old. This system might be beneficial for making upper body garments. Getting the right sized garments appropriate for individual body type and size based on the right size designation can be more comfortable and pleasant for the wearer.

REFERENCES

[1] Faust, M. E., & Carrier, S. (2010). Women's wear sizing: a new labelling system. *Journal of Fashion Marketing and Management: An International Journal*, 14(1), 88-126.  
 [2] Field, A. (2005). Factor analysis using SPSS. C8057 (Research Method II): Factor Analysis on SPSS.  
 [3] Goldsberry, E., Shim, S., & Reich, N. (1996). Women 55 years and older: Part II. Overall satisfaction and dissatisfaction with the fit of

ready-to-wear. *Clothing and Textiles Research Journal*, 14(2), 121-132.  
 [4] Gupta, D., & Gangadhar, B. R. (2004). A statistical model for developing body size charts for garments. *International Journal of Clothing Science and Technology*, 16(5), 458-469.  
 [5] <http://www.fibre2fashion.com/industry-article/4/393/indian-body-dimensions1.asp>  
 [6] Hsu, C. H., & Wang, M. J. J. (2005). Using decision tree-based data mining to establish a sizing system for the manufacture of garments. *The International Journal of Advanced Manufacturing Technology*, 26(5-6), 669-674.  
 [7] McCulloch, C. E., Paal, B., & Ashdown, S. P. (1998). An optimisation approach to apparel sizing. *Journal of the Operational Research Society*, 49(5), 492-499.  
 [8] Petrova, A. (2007). Creating sizing systems. *Sizing in clothing: Developing effective sizing systems for ready-to-wear clothing*, 57-87.  
 [9] Zakaria, N. (2011). Sizing system for functional clothing—Uniforms for school children. *Indian Journal of Fibre & Textile Research*, 36:348-357.