Chapter One
Introduction

1.1 Introduction

clustering is the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups .It is a main task is exploratory mining. Clustering methods are based on measuring distances between records and between clusters. Records are assigned to clusters in a way that tends to minimize the distance between records belonging to the same cluster. [3]

The main reason for having many clustering methods is the fact that the notion of “cluster” is not precisely defined. Consequently, many clustering methods have been developed, each of which has its own characteristics. [1]

The main goal of this research is to use clustering techniques to find out what extend the stander international sizing system poshirt Sudanese.

1.2 Problem Definitions

This research is found to answer the following question: “Is the standard sizing system is optimal for Sudanese men?”

In Sudan there has been no studies carried out in the field of sizing and fitting system.

The sizing systems used by clothing manufacturers for men's wear vary widely in size ranges and measurement. Most these measurements are very different from the standards recommended by the International Standards Organization (ISO).

1.3 Objectives

The objectives of this research are:

- To evaluate and define the sizing system.
- To explore the body types of Sudanese male based on the collected data set.
- To suggest new sizing for Sudanese (male).

1.4 Research Methodology

In this research, the Knowledge discovery in database (KDD) road map will be followed as described in [2]. To cluster the collected data which is sizing data, K-means will be used as selected clustering method.

KDD is a comprehensive Data Mining technique; therefore, steps irrelevant to our problem will be skipped.

Weka software will be used for performing the experiments of this thesis.

Chapter Two

Literature Review

2.1 Introduction

In the last twenty years there was an extraordinary expansion of computer accessible data about all kinds of human activities. The availability of these large volumes of data, and our limited capabilities to process them effectively, creates a strong need for new methodologies for extracting useful, task – oriented methodologies for deriving plausible knowledge from small and directly relevant data. However, in many practical areas only limited data may be available, e.g., fraud detection, terrorism presentation computer intrusion detection, early cancer diagnosis. [1]

In order to automatically generate useful knowledge from a variety of data, and presented it in human oriented forms, a powerful tools is strongly needed. Researchers have been exploring ideas and methods in different areas as efforts to satisfy this need. Such areas include; data mining, text mining, machine learning, statistical data analysis, data visualization, and pattern recognition. [2]

2.2 Data Mining

As mentioned earlier Data mining (DM, Data mining) is the nontrivial extraction of information that resides implicitly in the data. This information was previously unknown and may be useful for some process. In other words, prepares data mining, data probes and explores to remove the hidden information in them. On the other hand, the term knowledge discovery in databases (KDD) is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data (Fayyad, Piatetsky – Shapiro, and Smyth 1996). Here, data are a set of facts (forgeable, cases in a database), and pattern is an expression in some language describing a subset of the data or a model applicable to the subset. Hence, in our usage here extracting a pattern also designates fitting a model to data the term process implies that KDD comprises many steps, which involve data preparation, search for patterns, knowledge evaluation, and refinement, all repeated in multiple iterations. By nontrivial, we mean that some search or inference is involved; that is, it is not a straight forward computation of predefined quantities like computing the average value of a set of numbers. [2]

The discovered patterns should be valid on new data with some degree of certainty. We also want patterns to be novel (at least to the system and preferably to the user) and potentially useful, that is lead to some benefit to the user or task. Finally, the patterns should be understandable, if not immediately then after some post processing. [2]

The distinction between the KDD process and the data – mining step (within the process) is that. KDD refers to the overall process of discovering useful knowledge from data. Data mining (DM) is the application of specific algorithms for extracting patterns from data. Figure (1) shows the additional steps in the KDD process, such as data preparation, data selection, data cleaning, incorporation of appropriate prior knowledge, and proper interpretation of the results of mining, are essential to ensure that useful knowledge is derived from the data. [2]

2.3 Data Mining Functionalities

The kinds of pattern that can be discovered depend upon the data mining tasks employed. The main two types of data mining tasks are:

1**) Descriptive data mining tasks** that describe the general properties of the existing data. Descriptive data mining describes a dataset in concise way and presents interesting characteristics of the data without having any predefined target. [4]

2) **Predictive data mining tasks** that attempt to do predictions based on inference on available data. The term predictive data mining is usually applied to identify data mining projects with the goal to identify a statistical or neural network model or set of models that can be used to predict some response of interest. A predictive model is made up of a number of predictors, which are variable factors that are likely to influence future behavior or results. Figure (1): shows these two types of data mining tasks.[4]

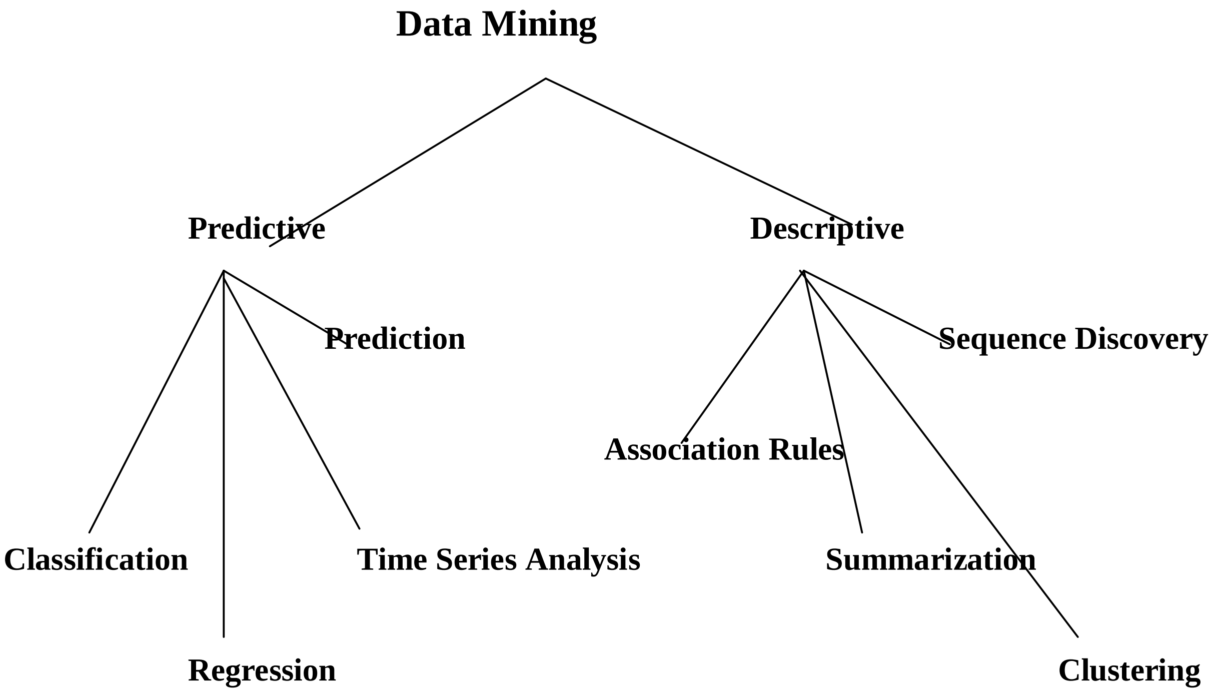


Figure 1: Data Mining Model and Tasks [4]

2.4 Clustering

Clustering is a division of data into groups of similar objects. Representing the Data by fewer clusters necessarily loses certain fine details, but achieves simplification. It models data by its clusters. Data modeling puts clustering in a Historical perspective rooted in mathematics, statistics, and numerical analysis. From a machine learning perspective clusters correspond to hidden patterns, the Search for clusters is unsupervised learning, and the resulting system represents a Data concept. From a practical perspective clustering plays an outstanding role in Data mining applications such as scientific data exploration, information retrieval And text mining, spatial database applications, Web analysis, CRM, marketing, Medical diagnostics, computational biology, and many others.[3]

Clustering is the subject of active research in several fields such as statistics, Pattern recognition and machine learning. This survey focuses on clustering in Data mining. Data mining adds to clustering the complications of very large Datasets with very many attributes of different types. This imposes unique Computational requirements on relevant clustering algorithms. A variety of Algorithms have recently emerged that meet these requirements and were successfully applied to real-life data mining problems. [3]

2.5 K-Means Clustering

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centrist, one for each cluster. These centrist should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centric. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centrist as bar centers of the clusters resulting from the previous step. After we have these k new centrist, a new binding has to be done between the same data set points and the nearest new centric. A loop has been generated. As a result of this loop we may notice that the k centrist change their location step by step until no more changes is done. In other words centrist do not move any more [3].

2.6 Determining the Software Package

To choose the most suitable software package is not an easy task since from the review of earlier research, software packages have their own strengths and weaknesses. Based on the information known about developing sizing system for poshirt (suit) for Sudanese men it was decided to select, weka 3.6.4 which has been implemented in Java with latest windows 7 operating in Intel Core 2Quad@2.83 GHz and 2GB memory, with a fairly simple macro processer as a software allows for the data to be imported into their software directly from Excel. This is an advantage because; all the data collected from anywhere can be downloaded into Excel, because it is well known and used program.

2.7 Sizing Issues

Size is (dependent) related to individual body dimensions that can vary during one person's lifetime and from one person to another, as well as from one generation to the next. However, time, sex and race, are not the only factors that affect sizing. Consumer education and behavior, along with often marketing issues have an impact on sizing too [5].

When compare between manufactures of items who employee their own fit models and size charts, it would be found that there are widespread differences in sizing and fit. These variations such as, use of garment, location of the garment, style fashion, fabric, ease allowance, and target selling price, all contribute to defining the size of a product. Consumer education and perception have a direct impact on sizing issues, Therefore they are related to a company's marketing strategy [6].

Chapter Three

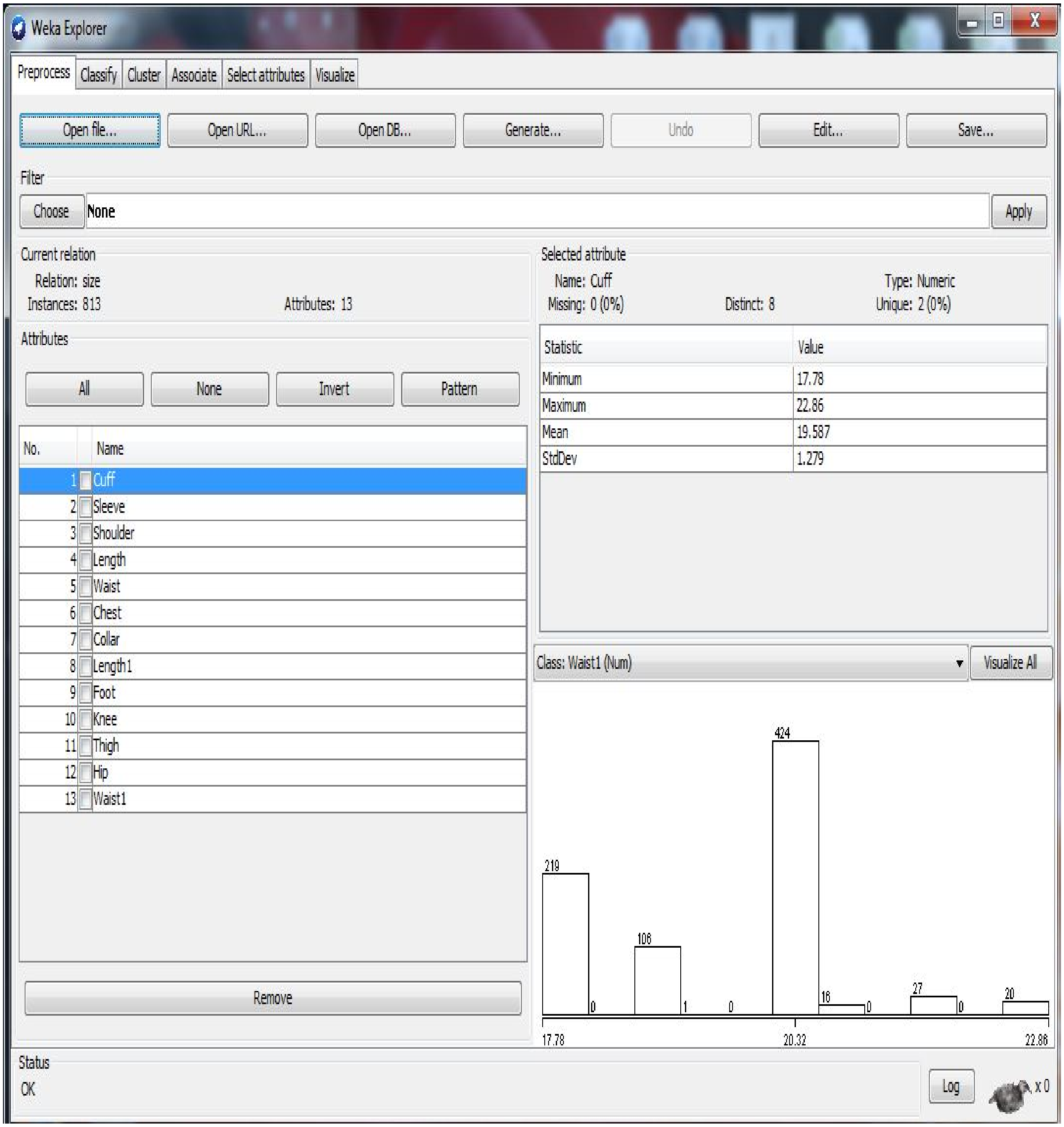
Methodology and Data

3.1 Introduction

This chapter introduces the data mining package used in these experiments which is Weka [7]. The second part of this chapter introduces the data set used in the study.

3.2 Weka (machine learning)

Weka (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is free software available under the GNU General Public License [7]. Show figure 2.



Filter

None

Current relation

Relation: size

Instances: 813 Attributes: 13

Selected attribute

Name: Cuff Type: Numeric

Missing: 0 (0%) Distinct: 8 Unique: 2 (0%)

Statistic	Value
Minimum	17.78
Maximum	22.86
Mean	19.587
StdDev	1.279

Attributes

No.		Name
1	□	Cuff
2	□	Sleeve
3	□	Shoulder
4	□	Length
5	□	Waist
6	□	Chest
7	□	Collar
8	□	Length1
9	□	Foot
10	□	Knee
11	□	Thigh
12	□	Hip
13	□	Waist1

Class: Waist1 (Num)

Status

OK

Fig 2: Weka version 3.6.9 main page

3.2.1 Description

The Weka (pronounced Weh-Kuh) contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access.

The original non-Java version of Weka was a TCL/TK front-end to (mostly third-party) modeling algorithms implemented in other programming languages, plus data preprocessing utilities in C, and a Make file-based system for running machine learning experiments. This original version was primarily designed as a tool for analyzing data from agric lateral domains,[2][3] but the more recent fully Java-based version (Weka 3), for which development started in 1997, is now used in many different application areas, in particular for educational purposes and research. Advantages of Weka include:

- free availability under the GNU General Public License
- portability, since it is fully implemented in the Java programming language and thus runs on almost any modern computing platform
- a comprehensive collection of data preprocessing and modeling techniques
- ease of use due to its graphical user interfaces

Weka supports several standard data mining tasks, more specifically, data preprocessing, clustering, classification, regression, visualization, and feature selection. All of Weka's techniques are predicated on the assumption that the data is available as a single flat file or relation, where each data point is described by a fixed number of attributes (normally, numeric or nominal attributes, but some other attribute types are also supported). Weka provides access to SQL databases using Java Database Connectivity and can process the result returned by a database query. It is not capable of multi-relational data mining, but there is separate software for converting a collection of linked database tables into a single table that is suitable for processing using Weka.[4] Another important area that is currently not covered by the algorithms included in the Weka distribution is sequence modeling.

Weka's main user interface is the Explorer, but essentially the same functionality can be accessed through the component-based Knowledge Flow interface and from the command line. There is also the Experimenter, which allows the systematic comparison of the predictive performance of Weka's machine learning algorithms on a collection of datasets.

The Explorer interface features several panels providing access to the main components of the workbench:

- The Preprocess panel has facilities for importing data from a database, a CSV file, etc., and for preprocessing this data using a so-called filtering algorithm. These filters can be used to transform the data (e.g., turning numeric attributes into discrete ones) and make it possible to delete instances and attributes according to specific criteria.
- The Classify panel enables the user to apply classification and regression algorithms (indiscriminately called classifiers in Weka) to the resulting dataset, to estimate the accuracy of the resulting predictive model, and to visualize erroneous predictions, ROC curves, etc., or the model itself (if the model is amenable to visualization like, e.g., a decision tree).
- The Associate panel provides access to association rule learners that attempt to identify all important interrelationships between attributes in the data.

- The Cluster panel gives access to the clustering techniques in Weka, e.g., the simple k-means algorithm. There is also an implementation of the expectation maximization algorithm for learning a mixture of normal distributions.
- The Select attributes panel provides algorithms for identifying the most predictive attributes in a dataset.
- The Visualize panel shows a scatter plot matrix, where individual scatter plots can be selected and enlarged, and analyzed further using various selection operators.

3.3 Clustering in Weka

WEKA contains “clusterers” for finding groups of similar instances in a dataset. The clustering schemes available in WEKA are *k*-Means, EM, Cobweb, *X*-means, and FarthestFirst. Clusters can be visualized and compared to “true” clusters (if given). Evaluation is based on log likelihood if clustering scheme produces a probability distribution [4].

3.4 Research Data Collection

With consulting domain experts anthropometric sizing variables that are necessary and strongly associated with the production of safari suit are selected. The selected variables are as follows; for the upper body dimensions (shirt); shoulder length, waist circumference, chest length, collar length, sleeve length, and the shirt length. However, for the lower body dimensions (trousers) we have the following measures: waist circumference, hip circumference, length, foot length and leg length. Anthropometric data of 813 people (male) from the SUR military clothing factory, with the age ranged from (25 – 66) years old are measured. These measurements of anthropometric data followed the ISO 8559/1989 body measurements standard. The eleven (11) dimensions for safari (suit) measured by tape measure for measuring either body contours or the length of the body curvatures. The equipment was calibrated with accuracy level up to 1 centimeter. These records have existing measurements authorized by anthropometric experts [5].

3.5 Data Cleaning

Before clustering the data, the following cleaning tasks were performed on the data:

- Fill Missing values
- Convert data to a standard format (csv)
- Fix errors and outliers
- Remove Noisy data

Chapter Four

Result and discussion

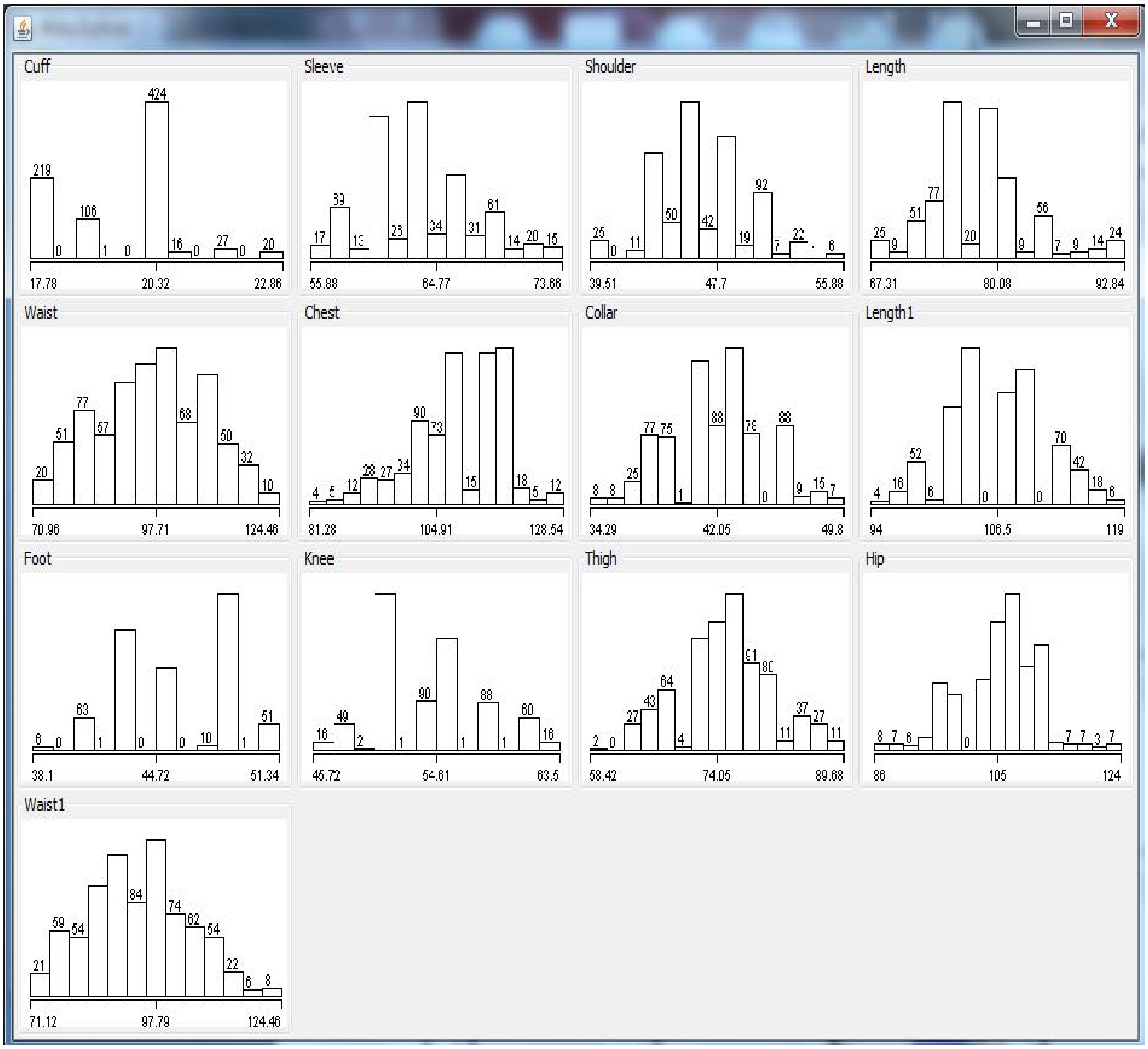
4.1 data set Description

Table 1, summarizes the data set used in thesis experiments. Figure 3, shows data attributes distribution.

Table 1

Data set characteristic:	Numeric	Number of instances:	813
Attribute characteristic:	Real	Number of attributes:	13
Associated tasks:	Clustering	Missing values:	No
Attribute names:	1- cuff	2- sleeve	3- shoulder
	4- length	5- waist	6- chest
	7- color	8- length	9- foot
	10- knee	11- thigh	12- hip
	13- waist		

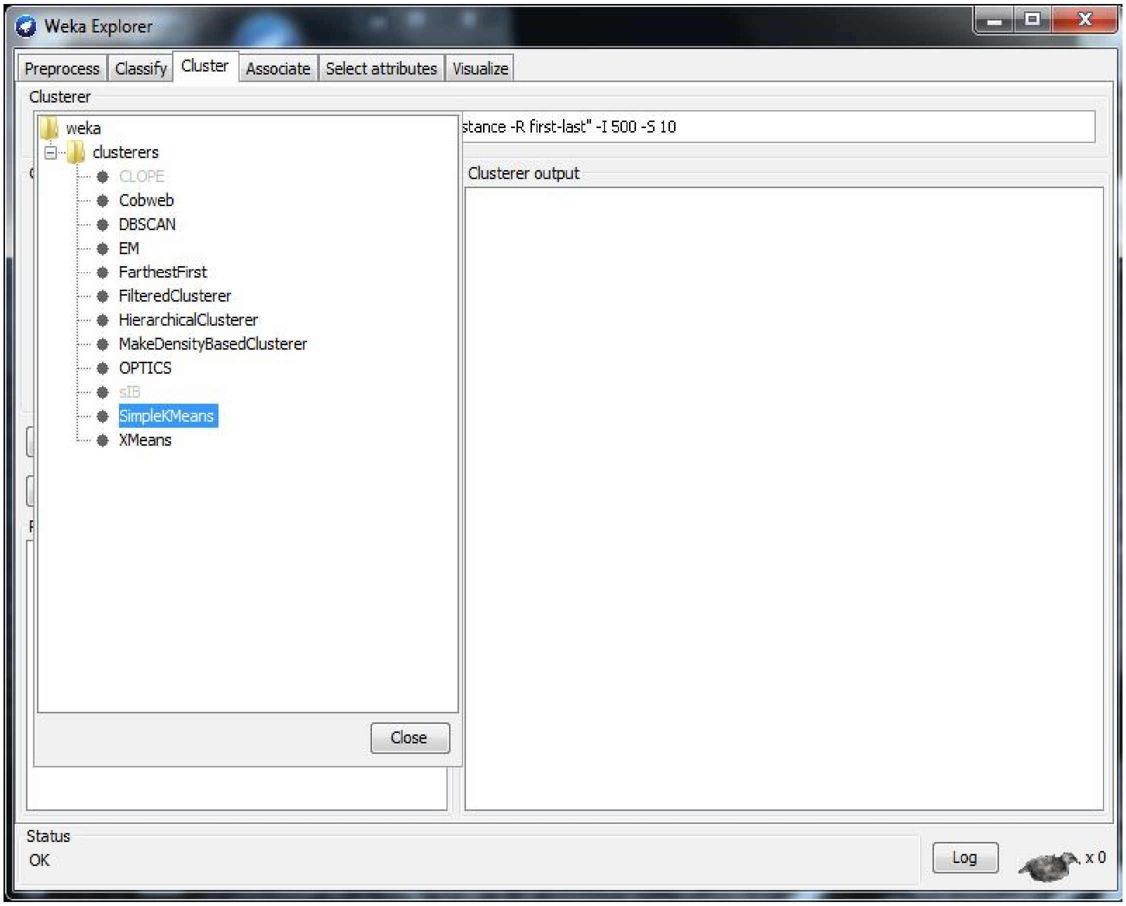
Figure 3: visualization for data set attributes distribution



4.2 Experiments description

Using the given data set three experiments were performed. The description of these experiments and their discussions is given in the following sections.

Fig 4: Snap shot for clustering page in weka

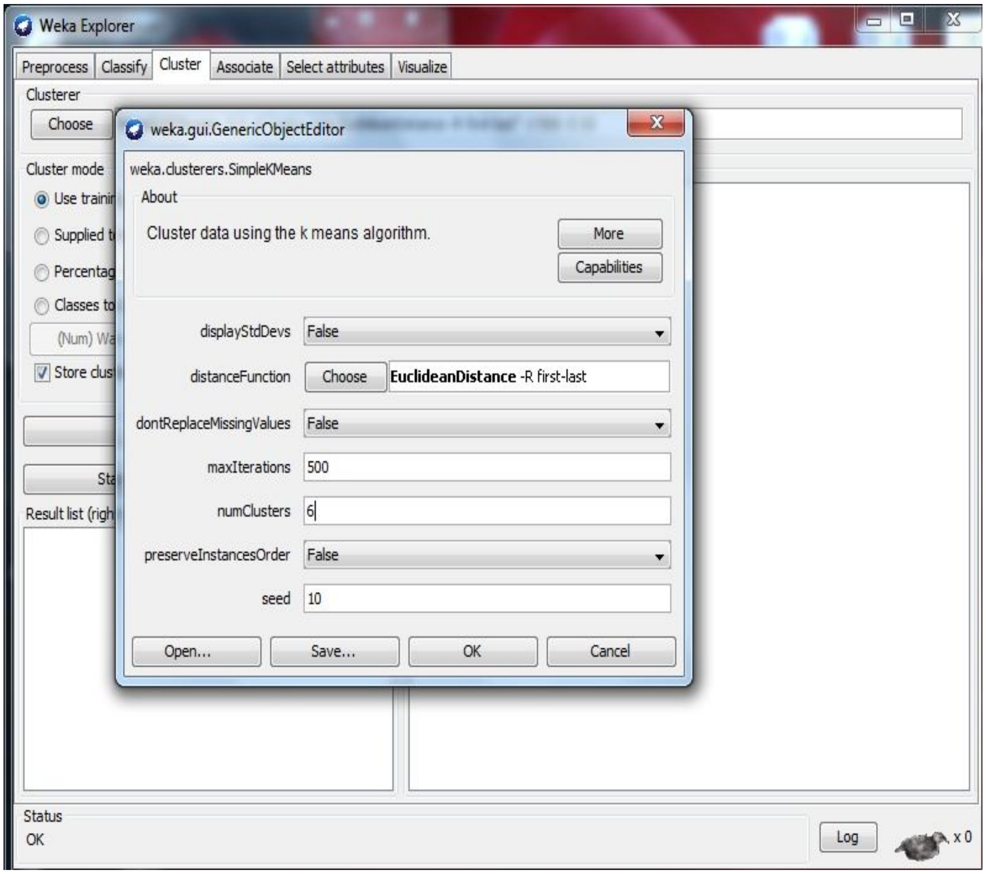


```
stance -R first-last" -I 500 -S 10
```

4.2.1 Experiment One

In this experiment we had used the weka to cluster our data set into 6 clusters. See figure 5, Table 2 contains the result of this experiment.

Fig 5: 6 clusters experiment



weka.clusterers.SimpleKMeans

Cluster data using the k means algorithm.

displayStdDevs: False

distanceFunction: **EuclideanDistance** -R first-last

dontReplaceMissingValues: False

maxIterations: 500

numClusters: 6

preserveInstancesOrder: False

seed: 10

Table 2: The table shows the found clusters centroids

Attribute	Full data (813)	Cluster 0 (130)	Cluster 1 (138)	Cluster 2 (134)	Cluster 3 (167)	Cluster 4 (94)	Cluster 5 (150)
Cuff	19.6	20.2	20.4	18.2	20.1	17.9	20.1
Sleeve	63.6	66.1	62.7	63.9	63.5	62.1	63.0
Shoulder	46.6	49.1	45.6	44.9	46.5	44.9	47.8
Length	78.2	83.4	75.9	76.1	77.4	74.4	80.9
Waist	96.7	109.1	89.6	88.7	97.1	86.7	105.3
Chest	109.1	114.9	105.7	106.7	108.1	105.5	112.3
Collar	41.9	44.2	41.1	39.6	41.6	40.1	44.3
Length1	106.3	108.4	105.7	106.7	106.6	104.2	105.6
Foot	45.8	48.5	44.3	47.5	47.7	42.9	43.1
knee	53.9	58.1	50.9	54.4	57.3	49.6	51.6
Thigh	75.2	80.6	71.2	74.2	76.7	68.8	77.2
Hip	105.2	105.5	105.4	105.4	104.5	104.7	105.5
Waist1	94.1	105.5	86.4	88.9	96.1	83.3	100.2

4.2.1.1 Result analysis

Cluster 0 resembles **2XL** which contains 130 records, about 16% of the data.

Cluster 1 resembles **S** which contains 138 records, about 17% of the data.

Cluster 2 resembles **M** which contains134 records, about 16% of the data.

Cluster 3 resembles **L** which contains 167 records, about 21% of the data.

Cluster 4 resembles **XS** which contains 94 records, about 12% of the data.

Cluster 5 resembles **XL** which contains 150 records, about 18% of the data.

Table 3: This table shows the international standard shirt size for the six standards together the nearest clusters shirt size (chest).

Size status	International standard	Data set	cluster	Percent%
2XL	116.8	114.9	0	16%
XL	114.3	112.3	5	18%
L	111.8	108.1	3	21%
M	109.2	106.7	2	16%
S	106.7	105.7	1	17%
XS	104.1	105.5	4	12%

Table 4: This table shows the international standard trousers size for the six standards together the nearest clusters trousers size (waist).

Size status	International standard	Data set	cluster	Percent%
2XL	109	105.5	0	16%
XL	104	100.1	5	18%
L	99	96.1	3	21%
M	94	88.9	2	16%
S	89	86.4	1	17%
XS	84	83.3	4	12%

4.2.1.2 Experiment conclusion

The first experiment shows that:

- No clusters resemble size 3XL.
- The average of difference between the found clusters and the standard ones is 2.64.

4.2.2 Experiments Two

In these experiments we increase the number of clusters to 7.

Table 5: The table shows the result of experiment two

Attribute	Full data (813)	Cluster 0 (107)	Cluster 1 (133)	Cluster 2 (100)	Cluster 3 (175)	Cluster 4 (76)	Cluster 5 (154)	Cluster 6 (68)
Cuff	19.6	20.4	20.4	18.5	19.9	17.8	20.1	18.2
Sleeve	63.6	65.7	62.4	63.2	64.3	61.1	62.4	67.2
Shoulder	46.6	49.6	45.7	44.6	46.5	44.9	48.1	45.3
Length	78.2	83.3	75.9	75.0	78.4	73.5	80.4	78.9
Waist	96.7	110.6	89.4	85.2	97.5	85.9	105.2	96.2
Chest	109.0	115.9	106.2	106.2	108.7	105.1	112.4	105.8
Collar	41.9	44.6	41.1	38.5	41.7	39.7	44.2	42.5
Length1	106.3	108.5	105.4	106.1	106.8	103.8	105.0	109.2
Foot	45.8	48.5	44.2	48.1	47.9	43.0	43.4	44.6
knee	53.9	58.2	50.9	54.7	57.6	49.6	51.9	52.1
Thigh	75.2	81.0	71.1	73.4	77.2	68.6	77.1	74.0
Hip	105.2	106	105.1	105.6	104.3	105.4	105.7	104.0
Waist1	94.0	106.8	85.9	86.0	96.9	81.9	100.2	93.7

4.2.2.1 Results Analysis

Clusters 0 resembles **2XL** which contains 107 records, about 13% of the data.

Clusters 1 resembles **S** which contains 133 records, about 16% of the data.

Clusters 2 resembles **S** which contains 100 records, about 12% of the data.

Clusters 3 resembles **M** which contains 175 records, about 22% of the data.

Clusters 4 resembles **XS** which contains 76 records, about 9% of the data.

Clusters 5 resembles **XL** which contains 154 records, about 19% of the data.

Clusters 6 resembles **S** which contains 68 records, about 8% of the data.

Table 6: This table shows the international standard shirt size for the seven standards together the nearest clusters shirt size (chest).

Size status	International standard	Data set	cluster	Percent%
2XL	116.8	115.8	0	13%
XL	114.3	112.3	5	19%
M	109.2	108.7	3	22%
S	106.7	106.2	2	12%
S	106.7	106.1	1	16%
S	106.7	105.7	6	8%
XS	104.1	105.0	4	9%

Table 7: This table shows the international standard trousers size for the seven standards together the nearest clusters trousers size (waist).

Size status	International standard	Data set	cluster	Percent%
2XL	109	106.8	0	13%
L	99	100.2	5	19%
L	99	96.9	3	22%
M	94	93.6	6	8%
XS	84	86.2	2	12%
XS	84	85.9	1	16%
XS	84	81.9	4	9%

4.2.2.2 Experiment conclusion

The second experiment shows that:

- Still no clusters resemble 3XL.
- In table 6there are 3 clusters resemble the standard size S, with 318 records and with total percentage 36% all the 3 clusters are less than the international standard S.
- In table 7 there are 3 clusters resemble the standard size XS, with 254 records and with total percentage 37% the 3 clusters are more than the international standard XS.

4.2.3 Experiment Three

In this experiment we increase the number of cluster to 8.

Table 8 show the result of this experiment

Attribute	Full data (813)	Cluster 0 (105)	Cluster 1 (75)	Cluster 2 (97)	Cluster 3 (161)	Cluster 4 (128)	Cluster 5 (64)	Cluster 6 (63)	Cluster 7 (120)
Cuff	19.6	20.4	17.9	18.5	20.1	20.4	18.7	18.0	20.6
Sleeve	63.6	65.7	62.1	63.1	64.1	62.5	59.4	67.6	63.7
Shoulder	46.6	49.6	44.5	44.7	46.4	45.6	47.9	45.9	47.7
Length	78.2	83.3	73.8	75.1	78.1	75.9	77.8	79.2	81.1
Waist	96.7	110.8	83.9	84.9	97.4	89.4	102.6	96.7	105.3
Chest	109.0	116.1	105.4	105.5	108.7	105.6	110.9	107.5	111.9
Collar	41.9	44.6	39.6	38.5	41.7	41.2	43.2	42.4	44.2
Length1	106.3	108.2	104.7	106.3	106.7	105.5	101.7	109.3	106.8
Foot	45.8	48.5	42.9	48.0	47.9	44.2	44.3	45.9	43.2
knee	53.9	58.3	49.4	54.7	57.7	50.9	52.1	53.4	51.8
Thigh	75.2	81.0	67.9	73.2	77.0	70.9	76.2	75.6	77.3
Hip	105.2	105.8	104.9	105.8	104.4	105.5	104.4	104.0	105.9
Waist1	94.0	106.9	81.1	86.0	96.6	85.8	97.1	94.9	100.5

4.2.3.1 Analyzing Results

Clusters 0 resembles **2XL** which contains 105 records, about 13% of the data.

Clusters 1 resembles X**S** which contains 75 records, about 9% of the data.

Clusters 2 resembles **S** which contains 97 records, about 12% of the data.

Clusters 3 resembles **M** which contains 161 records, about 20% of the data.

Clusters 4 resembles **S** which contains 128 records, about 16% of the data.

Clusters 5 resembles **L** which contains 64 records, about 8% of the data.

Clusters 6 resembles **S** which contains 63 records, about 8% of the data.

Clusters 7 resembles **L** which contains 120 records, about 15% of the data.

Table 9: This table shows the international standard shirt size for the eight standards together the nearest clusters shirt size (chest).

Size status	International standard	Data set	cluster	Percent%
2XL	116.8	116.1	0	13%
L	111.8	111.9	7	15%
L	111.8	110.9	5	8%
M	109.2	108.7	3	20%
S	106.7	107.5	6	8%
S	106.7	105.6	4	16%
XS	104.1	105.5	2	12%
XS	104.1	105.4	1	9%

Table 10: This table shows the international standard trousers size for the eight standards together the nearest clusters trousers size (waist).

Size status	International standard	Data set	cluster	Percent%
2XL	109	106.9	0	13%
L	99	100.4	7	15%
L	99	97.1	5	8%
L	99	96.5	3	20%
M	94	94.9	6	8%
S	89	87.8	4	16%
XS	84	86.0	2	12%
XS	84	81.1	1	9%

4.2.3.2 Experiment conclusion

The third experiment shows that:

- Still no clusters resemble 3XL.
- No clusters resemble size XL.
- In table 10 there are 3 clusters resemble the standard size L with 294 records and with total percentage 43% one of the clusters is more than the international standard L.

Chapter Five
Conclusion & future work

5.1 Conclusion

The purpose of this research is comparing international standard sizing system (XS, S, M, L, XL, 2XL, 3XL), with our data set. There are 3 experiments have been performed to find the distribution of Sudanese (male) sizes.

The results show that:

- The size 3XL does almost not exist among the data.
- The majority of Sudanese are in the S sizes (36% for shirt) and L sizes (43% for trouser).

5.2 Future Work

- Increase the data set to get more information’s.
- Using other clustering techniques in weka or other software to get more information.
- Get women’s data and analyses it.
- Comparing international standard sizing system with our data set (another attribute such as Hip, length, waist, etc...).

Reference

[1] Han, J. and Kamber, K., Data mining: *Concept and Techniques*. San Francisco: Morgan Kaufman Publisher (2001).

[2] Witten, E. Frank, Data Mining, *Practical Machine Learning Tools and Techniques with Java Implementation*, Morgan Kaufmann Publishers, 2000.

[3] AGGARWAL, C.C., PROCOPIUC, C., WOLF, J.L., YU, P.S., and PARK, J.S. 1999a. *Fast algorithms for projected clustering*. In Proceedings of the ACM SIGMOD Conference,61-72, Philadelphia, PA.

[4] R. Bagherzadeh, ATMT Research Institute, Amirkabir University of Technology, Tehran, Iran.E-mail: Bagherzadeh_r@aut.ac.ir, *World Appl. Sci. J., 8 (8): 923-929, 2010*

[5] Meng, J.C., L. Hai and J.J.W. Mao, 2007. *Thedevelopment of sizing systems for Taiwanese elementary- and high-school students*, International J. Industrial Ergonomics, 37: 707-716.

[6] Chang, C.F., 1999. *The model analysis of female bodythe size measurement* from 18 to 22, J. Hwa Gang Textile,6: 86-94.

[7] R. Kirkby, WEKA Explorer User Guide for version 3-3-4, University of Weikato, 2002.