**AUTOMATED MARKET BASKET ANALYSIS SYSTEM**

**ABSTRACT**

Due to technology Advancements, relationship marketing has become a reality in recent years. Technologies such as data warehousing, data mining, and campaign management software have made customer relationship management a new area where ﬁrms can gain a competitive advantage. Particularly through data mining the extraction of hidden and useful information from a large set databases companies can identify valuable customers, predict future behaviors based on the previous purchase pattern of the customer, and enable companies to make proactive, knowledge-driven decisions.

we propose a system that will help retailers to understand dependencies among goods purchased by the customer , also knowing what good is purchased with the other or if a particular set of goods are purchased so as to maximize profit.

# CHAPTER ONE

**1.0 INTRODUCTION**

Data mining is described as the extraction of hidden helpful information from a collection of huge databases, data mining is also a technique that encompasses an enormous form of applied mathematics and compultational techniques like link analysis,clustering, classification, summarizing knowledge , regression analysis and so on. data mining tools predict future trends and behaviors, permitting businesses to create knowledge-driven selections. The machine-driven, prospective analyses offered by data mining move on the far side the analyses of past events. data mining tools provides answer to business questions that were time consuming. They search databases for hidden patterns, finding useful information that is beyond the reach of specialists.

Data mining techniques is enforced speedily on existing package and hardware platforms to reinforce the worth of existing information resources, and might be integrated with new product and systems as they're brought. once enforced on high performance client/server or multiprocessing computers, data mining tools will analyze huge databases to provide answers to questions such as, ”What goods consumers tend to buy the most and goods that go along side with it”.

Coenen(2010) in his publication” Data Mining: Past, Present and Future” discussed the history of data mining can be dated as far back as late 80s when the term began to be used, at

least within the research community and diffrentiated it from sql.

Broadly data mining can be defined as as set of mechanisms and techniques, realised in software,

to extract hidden information from data. However,the word hidden in this definition is important;

By the early 1990s data mining was commonly recognised as a sub process within a larger process called Knowledge Discovery in Databases or KDD ,the most commonly used definition of KDD is that of Fayyad et al as “the nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data.’’ (Fayyad et al. 1996).

As such data mining should be viewed as the sub-process, within the overall KDD process, concerned with the discovery of hidden information". Other sub-processes that form part of the KDD process are data preparation (warehousing, data cleaning, pre-processing,and so on) and the analysis/visualisation of results. For may practical purposes KDD and data mining are seen as synonymous, but technically one is a sub-process of the other. The data that data mining techniques were originally directed at was tabular data and, giventhe processing power available at the time, computational eficiencywas of significant concern. As the amount of processing power generally available increased, processing became less of a concern and was replaced with a desire for accuracy and a desire to mine ever larger data collections. Today, in the context of tabular data, we have a well established range of data mining techniques available.

It is well within the capabilities of many commercial enterprises and researchers to mine tabular

data, using software such as Weka, on standard desktop machines. However,the amount of electronic data collected by all kinds of institutions and commercial enterprises,year on year, continues to grow and thus there is still a need for efective mechanisms to mineever larger data sets. The popularity of data mining increased significantly in the 1990s, notably with the establishment of a number of dedicated conferences; the ACM SIGKDD(special intrest group on knowledge discovery in data) annual conference in 1995, and the European PKDD(practice of knowledge discovery in databases) and the Pacific/Asia PAKDD(pacific asiaconference on knowledge discovery and data mining) conferences This increase in popularity can be attributed to advances in technology; the computer processing power and data storage capabilities available meant that the processing of large volumes of data using desktop machines was a realistic possibility. It became common place for commercial enterprises to maintain data in computer readable form, in most cases this was primarily tosupport commercial activities, the idea that this data could be mined often came second. The 1990s also saw the introduction of customer loyalty cards that allowed enterprises to record customer purchases, the resulting datacould then be mined to identify customer purchasing patterns. Data mining , is the method of looking into giant volumes of data for patterns using methods like classification, association rule mining, clustering, etc.. data mining is a topic that is related to topics like machine learning and pattern recognition. data mining techniques area unit the results of an extended process of analysis and products development.

I am in my final year. I was bright and brilliant, my family was optimistic in me; they thought so much of me, but I had a fault. What was my fault? I hated compiler construction. I struggled with calculations all my life. Though i have been lucky; I did well all the same. However, I had to write my final exam. I searched for all Compiler construction past question for each year, compared, and sorted them. Guess what I discovered! Over 35% of the questions were repetitions. I had hit the jackpot. I carefully and thoroughly checked through the answer page. Therefore, I kept on revising only the repeated questions. Well, I have a good grade to show for the Data Mining I performed.

There is huge amount of data available in Information Industry. This data is of no use until converted into useful information. Analyzing this huge amount of data and extracting useful information from it is necessary. The extraction of information is not the only process we need to perform; it also involves other processes such as Data pre-processing( Data Cleaning, Data Integration, Data Transformation) Data Mining, Pattern Evaluation and Data Presentation. Once all these processes are over, we are now position to use this information in many applications such as Fraud Detection, Market Analysis, Production Control, Science Exploration etc.

* 1. **PROBLEM STATEMENT**

Through in depth research and observations carried on supermarket we have discorvered that retailers are willing to know what product is purchased with the other or if a particular products are purchased together as a group of items. Which can help in their decision making with respect to placement of product , determining the timing and extent of promotions on product and also have a better understanding of customer purchasing habits by grouping customers with their transactions.

This project is aimed at designing and implementing a well-structured market basket analysis software tool to solve the problem stated above and compare the result to that of an existing software called WEKA.

* 1. **SIGNIFICANCE OF THE STUDY**
	2. **AIM AND OBJECTIVE OF THE STUDY**

The aim of the study is to maximize profit for the retailers by providing better services to the consumers

The objective of this study are:

* Cross-Market Analysis - Data Mining performs Association/correlations between product sales.
* Identifying Customer Requirements - helps in identifying the best products for differentcustomers. It uses prediction to find the factors that may attract new customers.
* Customer Profiling - helps to determine what kind of people buy what kind of products.
	1. **METHODOLOGY**
1. **Data Pre-Processing**

Due to the fact that the data we are getting is a raw data,raw data in the real world may be incomplete it has to be pre-processed the raw data has to go through **data cleaning,data integration,data normarlization,data reduction** because without a quality data there will be no quality mining results.

* **data cleaning:**This has to do withfilling of missing values, resolving of inconsistencies in the raw data.
* **data integration:**combining data from multiple sources and generating the user with unified view of the data
* **normarlization:** normalization is used to minimize or to reduce redundancy.
* **data reduction:**reduction of the data set that is much smaller in volume but yet yields the same analytical results

**1.5 SCOPE OF THE STUDY**

This scope of the study focuses on Babcock Ventures supermarket and the scope of this project includes:

1. We aim to develop our very own market basket analysis software, which will be used in babcock university
2. The software will exhibit a colorful GUI(graphical user interface).
3. The software will be based on Apriori .
4. We intend to conduct a research into the various branches of science that this software will be based on, such as artificial intelligence.
5. We will develop a software that will eventually stand out among other data mining software.

**1.6 LIMITATION OF THE STUDY**

The limitations of this software will include:

1. **Data restrictions:**this is a major factor that stands in the way of the execution of this project.Since there is no data on households and individual consumers ,we neglect such purchases.
2. **Time constraints:** this is also a major factor due to the fact that it can’t work on a small amount of raw data because it tends to mislead the retailer in a nut shell this software will work on large volumes of data.

# CHAPTER TWO

# LITERATURE REVIEW

# INTRODUCTION

Market basket analysis is a model of association rule mining, association rule mining is a rule under the descriptive aspect of data mining, association rule is used to analyze relationship between a large volume of data items such as discovering of hidden association between quantities like finding all items which are frequently purchased with tea. Association rule mining finds all rules in the database that satisfy some minimum support and minimum confidence constraints

 Agrawal and Srikant (1994). The target of mining is not predetermined in association rule mining



Punj and Stewart (1983) reviewed the applications of cluster analysis in promoting issues and that suggested a two stage cluster analysis methodology of clusters through Ward’s minimum variance method. Theymentioned the issues and problems associated with the employment and validation of cluster analysis. Agrawal et al. (1993), in their study, delineate that within the recent past the exploratory analysis especially that of enormous sets of market basket knowledge has become topic of pertinent analysis because of varied publications on data processing and information in databases and generated association rules from market basket data, that describe relevant interdependencies.Piatetsky-Shapiro et al. (1996) surveyed a growing range of commercial applications of data mining. They have examined the prevailing data mining tools, delineate some representative applications like promoting, investment, producing, fraud detection etc. and mentioned the problems for deploying made application and their adoption by business users.

 Spiller and Lohse (1997), in their paper, gift a classification of on-line retail stores based mostly upon convenience sample of 137 internet retail stores. Cluster and correlational analysis known five distinct internet catalog interface classes which offer a far better understanding of the ways pursued in Internet-based promoting . Collier et al. (1998) mere that in the Nineties several firms reached to the conclusion that their data could be a valuable quality.

These firms sped to created datastore and data Marts. moreover, firms like Wal-Mart recognized the advantage of applying data mining to those wealthy stores of historical data .Weiss and Indurkhya (1998) explicit that a business services company with national reach asked data Miners to assist them improve their resource allocation by understanding the characteristics of the places wherever they were most successful. Michel and John (1998) centered on fourteen data mining tools and so a standardize procedure and twenty analysis criteria for assessing the tool qualities were developedand applied. The traits were collected in 5 categories; capability, learn ability, Stability, ability, Flexibility and Accuracy. They then summarized the analysis procedure and also the evaluation of all part criteria .Koga (1998) delineate that data mining techniques will discover patterns. Koga suggested that techniques in data mining, like association, classification, and clump are helpful for locating the mixture of product that have high chances of purchase at identical time, finding dynamical patterns of demand, distinguishing non-loyal customers, and targeting customers to extend response rate of information mining . Menon and Sharda (1999), stated that decision tree helps to classify information into a finite range of categories by generating a hierarchy of ‘IF-THEN’ statements. supported a series of ‘IF-THEN’ statements, they might predict the searching and patronage behaviors of retail customers.Witten and Frank (2000), in their study, discussed that data mining is that the task of distinctive helpful regularities and patterns in massive volume of data. the task involves distinctive patterns of client behavior from buying logs .

Russell and Petersen (2000), market basket analysis focuses on the choice method by that a client selects things from a given set of product classes on identical searching trip. It aims at the identification of interrelations between decisions of various merchandise purchased in a very specific outlet like a grocery Ma et al. (2000), data mining is important to the enterprise that wishes to use operational and different obtainable information to boost the standard of higher cognitive process and gain important competitive blessings. Hopping (2000) emphasized that the evolution of merchandising reveals that technology has contend a job because the primary enabler of amendment.Chopoorian (2001) explained that over the past years organizations have exerted vast efforts to making sure that they capture and store the maximum amount information concerning their business transactions and customers as potential. Today, a lot of and a lot of organizations area unit focusing their info technology efforts towards creating sense of that information. The authors any expand on the on top of statement by stating that information is that the structure challenge of the new millennium. They conjointly state that competitive success can depend upon the power of corporations to quickly and effectively convert their data into explicable info. it's not possible for managers to create the proper selections if the required info can not be accessed or conferred to them clearly.

 Song et al. (2001) delineated a way to notice changes of client behavior at totally different time from client profiles and sales information . Buck (2001), the retail analysis and money establishments are among the primary to embrace data processing technologies, initial with the analysis within the massive company data warehouses, and a lot of recently within the analysis of on-line web-based activities .

Bounsaythip and Rinta-Runsala (2001) provided a review on some data mining ways that may be used for client segmentation and identification.Katraras et al. (2001) used cluster analysis data processing technique to phase grocery shoppers into six segments in step with their preferences for thirty three store and searching expertise characteristics. The segmentation was done on the premise of angle and astonishingly as a result there have been only a few variations within the segments in terms of demographic factors. The six teams known by cluster analysis were Discriminating Leisure Shoppers, Time ironed meat eaters, Back to nature shoppers, Middle of the road. shoppers, one stop socialites and No nonsense shoppers. The characteristics rated most extremely by all the patrons were cleanliness and sanitation, followed by recent fruits and vegetables and also the quality of recent meat . archangel et al. (2001) concentrate on the very fact that a lot of of the helpful selling insights into client characteristics and their buying patterns area unit for the most part hidden and untapped. so a scientific methodology that uses data processing and data management techniques is projected to manage the selling data and is that the basis for enhancing client relationship management.. Lavinson (2002) specifies that retailers have collected large amounts of information for years, however they need not had the means that to use it effectively to their designing and shopping for as a result of, till a number of years agone, no laptop or code application may method all of the dataincorporated over simply historical sales information .

 Kopanas et al. (2002) mentioned that data mining techniques are applied in several application areas. a mining project has been typically delineated as a method of automatic discovery of recent knowledge from a massive quantity of information. The author has delineated data mining as a nonstop interaction between the implicit domain data and therefore the data that's discovered through the utilization of mining algorithmic rule . Nemati and Barko (2002) explicit that data is quickly changing into one among the foremost differentiators between business leading organizations and ordinary organizations. having the ability to extract the relevant information from this data plays a significant role in enhancing enterprise higher cognitive process. data mining may be a key a part of this technique, and its fortunate implementation will result in increased structure

 higher cognitive process .Coskun et al. (2002), data mining contains a generic blanket definition that tends to incorporate all the tools utilized to assist users analyze and perceive their information. They conjointly highlight that data data mining differs from ancient applied mathematics techniques by victimisation the pc, instead of the analyst, to seek out patterns and relationships by distinguishing the underlying rules and options within the information. The relationships ar thus found.inductively by the computer code utilized supported the prevailing information instead of requiring the creator to specify the purposeful type and interactions . Chris et al. (2002) offer the summary of the thought of information mining and CRM. The authors provide a more in-depth inspect in the main 2 data data mining techniques viz. Chi-Square Automatic Interaction Detection (CHAID) and Neural Networks. As a results of comparison of 2 techniques, the authors complete that CHAID is way easier and faster to construct and perceive whereas neural networks give additional correct models, particularly for complicated issues . Jackson (2002) provides an outline of data mining method. The author conjointly elaborates the thanks to set up, appraise and with success refine an information mining project, notably in terms of model building and analysis. The author concluded with a serious illustration of the mining method methodology and therefore the unresolved issues that supply opportunities for analysis. Elovici and Braha (2003) planned a call supposititious frame work for evaluating the information mining systems, that use classification strategies in terms of their utility in higher cognitive process. The model provides and economic perspective on the worth of extracted data in terms of its payoff to the organization, and counsel a good vary of call issues that will arise from this time of read. The author planned 2 means that by that freelance system will be combined that showed that the combined approach to data processing system will be utilized in the choice creating process of the organization to extend payoff . Labovitz (2003) explicit that data processing aims to feature worth to business organizations by applying extremely rigorous applied mathematics analysis to massive lots of information in an attempt to extract meaningful trends and relationships that ar typically hidden by the sheer volume and arrangement of the information. It any supports the transformation into information, data and knowledge. Mont and Plepys (2003) examined a good style of strategies for understanding and evaluating the consumer’s acceptance and satisfaction in numerous disciplines.

 Rosset and von Neumann (2003) mentioned in their paper, each in theory and through empirical observation, the optimum use of “customer value” altogether phases of information analysis like model coaching, model analysis and marking stages. Jaesung (2003) described, in his study, that the conception of data mining is gaining acceptance in business as a way of seeking higher profits and lower prices. Carrier and Povel (2003), the data mining consists of the building of the model from data. every data data mining technique will perform one or a lot of of the subsequent sorts of knowledge modeling like Association, Classification, Clustering, foretelling, Regression, Sequence Discovery,Visualization. The authors specialise in the purpose that the selection of knowledge mining techniques ought to be supported the information characteristics and business necessities . Lynette (2003) discusses a number of the common knowledge analysis techniques accustomed establish the foremost and least profitable customers in order that promoting managers will develop effective, suitably targeted client management ways . Levy and Weitz (2004) state that data mining assists organizations in making worth by providing functions like foretelling, modeling and support for deciding. RFM (recency, frequency, monetary) analysis utilized by catalog retailers and direct marketers, could be a theme for segmenting customers per however recent they need created a sale, however frequent they create purchases, and the way abundant they need bought . Shimizu (2004) examined however FSP (Frequent Shopper Program) knowledge ought to contribute to the retailers’ promoting strategy by victimisation data mining techniques. He used the FSP knowledge derived from a distributer for three months in 2002 and sampled 1296 transactions among the information hierarchical by quantity of purchase. He showed the steps to spot and classify customers supported real knowledge with data processing techniques. He advised a method of characteristic loyal customers, grouping loyal customers, assessing the profit of every loyal customers phase, ranking the segments, and checking ranking of segmentsWhile attention to data processing techniques has hyperbolic.

 Abe (2004) has questioned the recognition of knowledge mining techniques, that is classified as computer-based discovery of rules. Instead, he developed Associate in Nursing analytical model supported client behavior theories on ancient RF (recency and frequency) knowledge analysis and calculable the chance of the unobserved defection of shoppers .

 Tsai and Chiu (2004) stated that market segmentation is essential for an honest promoting and client relationship management program that is mostly done victimisation general variables like client demographic and fashion. that the author during this paper develops a unique market segmentation methodology supported product specific variables. conjointly purchase primarily based similarity live, bunch algorithmic program, and bunch quality operate area unit delineate during this paper. once finishing segmentation, a delegated RFM model is employed to investigate the relative profit of every client cluster Liu and Nilotic (2005) studied the implementation side of bunch data mining technique to client analysis of retail store. The study by Chen et al. (2005) integrates client activity variables, demographic variables, and dealing information to determine a way of mining changes in client. Kasindra and Robert (2005) demonstrated, using a test dataset, how a data mining process canhelp to achieve an integrated understanding of consumers by marrying information of various types and from various sources, in a manner that ensures that the resulting segments are logically and strongly differentiated on all the types of information in the analysis. Taking a case study from the telecommunications industry the authors produced behavioral and attitudinal segments in addition to the targeted segments. Bart et al. (2005) emphasized upon valuable relationship within the existing customers of a firm in the era of CRM. The authors have made the analyses on real-life sample of 100,000 customers taken from a data warehouse and then two types of random forests techniques were applied to analyze the data. Random forests are used for binary classification and regression forests for the model with linear dependent variables. Their findings suggested that past customer behavior is more important to generate repeat purchasing and favorable profitability evolutions while the intermediary’s role has a great impact on the customer’s defection proneness.

Chad et al. (2005) specify that the loyalty of customers to a super market can be measured in a variety of ways. Regular visitors and spenders are most likely to be loyal to a supermarket. The authors describe the results of experiments attempted to identify customer loyalty based on transactional data obtained from a supermarket data collection program. Wencai and Yu (2005) study the implementation aspects of applying clustering data mining method to customer analysis of a department store.Bhasin (2006) explains the importance of data mining tools in extracting important information from existing data to enable better decision making throughout the banking and retail industries.Data mining typically involves the use of predictive modeling, forecasting and descriptive modeling techniques.Wang and Hong (2006) described that the changes in customer behavior results in unpredictable customer profitability and cause inefficient and ineffective marketing planning. Wang and Wang (2006) explained that the online purchasing behavior is characterized by purchasing sequences. Their study proposes a data mining method for customer segmentation and applies it to an online nutrition product store. The results indicate that the method is novel and effective for the online customer segmentation . Jeroen et al. (2006) focused on the fact that the data mining methods like clustering and predictive analysis enable police force to get a clearer picture of criminal careers. Liu et al. (2007), in their study, emphasized on practical applications of association rulesin marketing stratagem to help rational marketing mix in retail enterprises . Sadic and Kayakutlu (2007), in their study, illustrated the implementation of cognitive maps and decision trees in the development of customer segments to be used by sales and marketing departments. Their study contributed not only in the field of data mining but also in customer relations . Millette (2007) discusses the need of a firm to determine ‘who are the best prospects among the existing customers for a new product’. For this regression, decision tree andneural network models are built to use for scoring the prospective customers. A confusion matrix is then used to determine the cutoff point for the scores for customers to determine who qualifies as a good target. The models are compared by assessing model performance and validating the model to new data . Buckinx et al. (2007) enriched the customer database with a prediction of customer’s behavioral loyalty so as to develop it for target marketing . Horng-Jinh et al. (2007) proposed an anticipation model of potential customers’ purchasing behavior, inferred from past purchasing behavior of loyal customers and the web server log files of loyal and potential customers by means of cluster and association analysis. Sung (2007), proposes a segment-based knowledge discovery method that can be used to derive the descriptive patterns such as individual transition paths and to predict the net path to which each customer is likely to shift. Hongwei et al. (2007) discussed that the organizations have accumulate a huge amount ofdata with the development in technology. The usage of data mining to tap the potentialknowledge is really crucial for retailers in the changing competitive environment. Shu-Hsien et al. (2008) presented a case study on the development and implementation of a data mining system in order to mine customer knowledge for electronic catalog marketing.

Yu and Wang (2008) proposed a hybrid data mining approach for e-retailers to analyzereturn patterns from both the customer and product perspectives. In the first stage theauthors use K-means algorithm to classify the customers and products according to theirdemographic data to predict their return patterns. Henry and Beverley (2008) present that the Knowledge Discovery in databases in a field of research that studies the development and use of various data analysis tools and techniques and Data mining is one of such tool. However, nearly two third of the ITmanagers say that data mining products are too difficult to use in a business context. Yasemin and Reutterer (2008) proposed and illustrated a two stage procedure combiningthe features form exploratory and model based approaches of market basket analysis.Retail marketing managers can make use of this information and thus can be assisted indesigning targeted direct marketing actions within their loyalty programs .Hsieh and Chu (2009) propose an integrated data mining and behavioral scoring model to manage existing credit card customers in a bank. A self-organizing neural network was used to identify groups of customers based on repayment behavior and RFM behavioral scoring predicators. Popovic and Dalbela (2009) aim to show that the data mining methods based on fuzzy logic could be successfully applied in retail banking analysis and they have usedapplication of fuzzy clustering in churn prediction for retail banking. Fang and Lee (2009) have collected 565 valid responses of the questionnaire distributedand then using two step cluster analysis, four distinct food-related consumer life stylesegments were identified and the segments differ in their attitude and behavior towardsfood consumption. Further the profiles of the segments are achieved by observing thesocio-demographic characteristics of typical segment members.Sheu et al. (2009) contributed to integrate data mining and experiential marketing to segment online game customers.Irko et al. (2009) describe that the customer segmentation is typically done by applying some form of cluster analysis to obtain a set of segments to which future customers areassigned to.Dhanpal et al. (2010) presented an original methodological approach of customer satisfaction evaluation by combining multicriteria preference disaggregation analysis and rule induction data mining. The authors states that using the customer characteristics,such as age, marital status, the presented methodology may identify and analyze special group of customers .Romdhane et al. (2010) describe that data mining is a new emerging discipline that aims at extracting knowledge from data using several techniques. Data mining proved to be useful in business where transactional data turned out to be a mine of information about customer purchase habits. Therefore developing customer models is an important step for targeted marketing. So the authors develop an approach for customer profiling composed of three steps. In the first step, the data was clustered in order to extract natural group of customers. Then the authors reduced the number of attributes in each group of customers using information entropy for the importance of an attribute. Finally neural network based modeling is used to built a set of customer profiles with the data in the corresponding group of customers .

## **BACKGROUND OF THE PROBLEM AREA**

Onifade(2013) in his presentation”Data Mining” discussed the need for data to minned.Over the last two decades companies have been intrested in understanding consumers buying habits and getting to know consumers better, the application of data mining techniques is that we tend to find useful information that can help organization make more profitable business decisions. Due to the vast deployment of Machine learning in data mining market analyst are no longer considered as the only factor trusted when it comes down to decision-making in business. Leone and Mulhern (1991) analyzed the impact of price promotions on cake mix and cake frosting. Their main objective is to evaluate the overall profitability of implicit price bundling. Reducing the price of cake mix increase purchases of both cake mix and frosting and the overall profit improves. The study shows how promotions have positive impact on the sales of a complementary product. Retailers wants to know how to arrange their product in order to attract customers,retailers have the oppurtunity of placing product that are sold together at diffrent segment of the supermarket so as to display other product that may as well intrest the customer or place product that are bought together side by side.This is an important reason why data needs to be minned so as to improve consumers satisfaction. Web merchants get the same benefit, by conveniently organizing their Web site so that items that sell together are found together. Outside of the store environment,market basket analysis provides different benefits, though equally useful ones. For a direct marketer, it is far preferable to market to existing customers, which are known to buy products and have a history with the company. The company already has these people in its database, and knows a significant amount of information about them. After running amarket basket analysis, a marketer can contact its prior customers with information about new products that have been shown to sell well with the products they've already bought; chances are, they'll be interested. In addition, even when making sales to new customers, telephone representatives can offer buyers of a product discounts on any other products they know sell with it, in order to increase the size of the sale.

## **APRIORI ALGORITHM**

There are various algorithms for Association Rule Mining that have been developed by various researchers to help users achieve their goals. Rakesh Agrawal and Usama Fayyad are one of the major pioneers in data mining. They account for a number of developed algorithms and procedures among which are apriori algorithm and aprioriTID algorithm,These algorithms are similar with regard to the function that is used to determine the itemsets, but the difference is that Apriori algorithm makes several passes over the database . The results from the study show that these two new algorithms perform much better than their predecessors known AIS algorithm developed by Agrawal, Imielinski, and Swami, 1993 and SETM algorithm by Houtsma and Swami, 1993. Since the introduction of the Apriori algorithm, it has been generally accepted as the most useful and fast algorithm for finding frequent itemsets.

Algorithm:

L1 = {frequent items};

**for** (k = 1; Lk !=∅; k++)

**do begin**

Ck +1 = candidates generated from Lk;

**for each** transaction t in database increment the count of all candidates in Ck+1that are contained in t

Lk +1 = candidates in Ck +1 with min\_support

**end**

**return** L = ∪k Lk;

Ck:Candidate itemset of size k Lk : frequent itemset of size k

Ck is generated by joining Lk-1 with itself

Any (k-1)-itemset that is not frequent cannot be a subset of a frequent k-itemset

## **2.4 ASSOCIATION RULE**

Association rule discovers interesting relationships among large volume of data items.

Let I is a set of items{i1,i2,….,im}.Let D is a set of transactions such that T I. Each transaction is uniquelyidentified with with an identifier called TID. The method can be stated as if there aretwo subsets of product items X and Y then an association rule is in the form of X→Ywhere X I and Y I. It implies that if a customer purchases X, then he or she also purchases Y. Two measures which reflect certainity of discovered association rules aresupport and confidence. Support measures how many times the transactional record in

database contain both X and Y. Confidence measures the accuracy of rule. For example, the information that customers who purchase diapers also tend to buy beer at the same time is represented in Association Rule below.

Diapers = Beer

Support = 30%, Confidence = 70%

Association rules are considered useful if they satisfy both a type equation here minimum support threshold and a minimum confidence threshold that can be set by users or domain consultants.

## **MARKET BASKET ANALYSIS**

## In order to perform market basket analysis, it is necessary to first have a list of transactions and what was purchased in each one. For simplicity, we will look at the example of purchase made by customers, each of whom bought only a few items:

Customer 1: Indomie,baked beans, vimto

Customer 2: Vimto, bread

Customer 3: Baked beans, indomie

Customer 4: Vimto, pie

Customer 5: Baked beans, pie

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | Indomie | Vimto | Baked Beans | Bread | Pie |
| Indomie | 2 | 1 | 2 | 0 | 0 |
| Vimto | 1 | 3 | 1 | 1 | 1 |
| Baked Beans | 2 | 1 | 3 | 0 | 1 |
| Bread | 0 | 1 | 0 | 1 | 0 |
| Pie | 0 | 1 | 1 | 0 | 2 |

Market basket analysis (MBA) is a data mining technique to discover associations between datasets. These associations can be represented in form of association rules.

**2.6 Advantages of market basket analysis**

1. Predict future trends, customer purchase habits
2. Help with decision making
3. Improve company revenue and lower costs

**2.7 Disadvantages market basket analysis**

1. User privacy/security
2. Amount of data is overwhelming

**2.8 REVIEW OF RELATED WORKS**

1. WEKA:The Waikato Environment for Knowledge Analysis is a widely toolkit data mining for and machine learning it was developed by a university in New Zealand called University of Waikato it entails a large collection of data mining algorithm written in java language. Weka contains tools for data preprocessing,visualisation,classification,clustering,association rule.it is said to be one of the most popular opensource data mining tools used for academic and research purposes
2. KNIME:In terms of a tool being user friendly, the Konstanz Information Miner is an open-source data integration,exploration,analysis and processing platform,it was bu developed using rigorous software engineering practices as blue prints KNIME is a modular data exploration platform that grant the user the ability to create pipelines(creating data flows).Developed by University of Konstanz
3. TANAGRA:This open source tool allows user io implement their own data mining method and to compare their performance ,it is said to be considered as a pedagogical tool for learning programming techniques as well also it offers several data mining methods statistical learning and machine learning and data analysis its primary aim is to provide a simplified version of data mining tool that is easy to use by the student and researchers.
4. ORANGE: Orange was developed using the C++ language it also has large variety of machine learning and data mining algorithms like WEKA but in addition to that some specific features were added such as as pretty-print of decision trees, attribute subset, bagging and boosting, and alike. Orange also includes a set of graphical widgets that use methods from core library and Orange modules.users can also develop and test their algorithm.it was developed at University of Ljubljana, inSlovenia
5. RapidMiner : The tool was developed using java it has user-friendly GUI. Everything in
	1. RapidMiner is focused on processes that may contain subprocesses like a nested if statment concept. What makes RapidMiner more intresting is the concept of the drag and drop operators which is used to construct the data flow.
6. SCIKIT-LEARN:scikit-learn is a free tool in Python that extends the functionality of NumPy and SciPy packages withvarious data mining algorithms. It also uses the matplotlib package for plotting charts. The package keeps improving by accepting valuable contributions from many contributors and is supported by both INRIA and Google Summer of Code.
7. R:The open-source tool and programming language that is mostly used for statisticians, R, is also a strong option fordata mining tasks. R has been in development for more than 15 years and is the successor of S, a statistical language originally developed by Bell Labs in 1970s. R is written in C++, Fortran, and in R itself.
8. XLMiner:This is an add-in for microsoft excel it performs comprehensive data mining likeneural nets,regression trees,it is very good in evaluation of predictive power
9. Rattle GUI:Its a graphical user interface developed by Dr.Graham Wiliams using the R statistical Language.it transforms data into forms that can br readily modelled.
10. ScaViS:developed in java and jython by Dr. Chekanov it means Scientific Computation And Visualisation Enviroment its is mainly used by scientist,engineers and students.

# CHAPTER THREE

# SYSTEM DESIGN/ DESIGN METHODOLOGY

## 3.0 INTRODUCTION

Analysis and Design phase of Software Development Life Cycle are very important stages and they involve careful observation and study of user requirements and existing system in comparison to the objectives of the proposed system. Thought the years systems, have evolved from a combination of manual components to a combination of software components such as modules, functions, classes, packages and applications. These software components are usually intangible replica of corresponding manual components. The components work together, sharing data, passing variables and processing data to give a final product that is geared towards the achievement of a particular goal, objective or set of objectives. Since design is key to the efficacy of delivery of user requirements, therefore on the long-run determines the validity of the system; this relationship translates directly to high costs of development reimplementation, and redesign if analysis and design phases are improperly organized.

## 3.1 SYSTEM ANALYSIS

System analysis is a problem-solving technique that decomposes a system into its component parts for the purpose of studying how well the component parts work and interact to accomplish their purpose. It describes what a system should do to meet the information needs of users.

System analysis is the process of enumerating the investigation of the existing system; problems of the existing system; analysing the proposed system for its costs and benefits; analysing both the system and user requirements and considering possible alternative systems. It is an in-depth study of end user information needs that produce functional requirements that are used as the basis for the design of a new information system.

Presumably, we do a system analysis in order to subsequently perform a systems design. (Whitten, Bentley and Dittman, 2001)

 **3.2 WATERFALL MODEL**

 The design that would be preferred for the development of this project is the waterfall model and this is primarily because the model prescribes a systematic approach to software development which starts with a well defined , understood specification of requirements and moves through to deployment in a linear form. The processes of arriving at this objective include:

* Requirement Analysis: This basically involves requirements gathering which requires a great deal of interaction/ communication and collaboration between the customers and software developers.
* Planning: This is the area where the overall plan for the engineering work that follows is developed and established. This plan involves the technical tasks and tools or resources, the likely risk and the work schedule for software project.
* Analysis and Design: Here we will be creating models for proper understanding of the system requirements and how best to achieve the requirements.
* Implementation: At this stage we will be working on the coding part of the software and after which we will run a test on the system to uncover possible errors.
* Deployment: The software is delivered as a complete product or in partial increment to the user who will evaluate and provide feedback on the evaluation.



## 3.3 PROPOSED SYSTEM DESIGN

As a result of the problems encountered in the existing system of operations in taking care of raw data, an improved system is developed to eradicate the shortcoming of the existing system.

The proposed system is being designed in such a way that in order to perform market basket analysis, the user must register for security reasons and the user can then upload raw data and then applies for the result of the analysis of the data.

# 3.4 PROGRAMMING LANGUAGE

Java is measured as the greatest application expansion language. Java is a general-purpose computer programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere" (WORA), meaning that code that runs on one platform does not need to be recompiled to run on another. It is an item-oriented programming language whichever is used to generate efficient quality apps for both mobile phones and computers. This was first presented in 1995 by James Gosling of Sun Micron Systems. This language has been established after counting the concepts of numerous languages similar to C, C++, etc. This is a flexible and platform liberated language.

## BENEFITS OF JAVA:

1. Simplicity: Java offers a very easy edge for the developers and users. This has been measured as the greenest language whenever compared to other software design languages. Java has removed the usage of pointers and also changed the difficulty of multiple traditions in C++ with an unassuming structure and that construction is called interface.
2. Platform and Portability Independent Conduct: Java is platform free. This provides the capability to Run Anywhere and Write Once. The applications advanced by using the language could be run on any software and hardware platform. The Java apps are held by each Java companionable browser.
3. Allocation: Java has feature of Stack provision system. This helps the statistics to be kept and could be re-established easily. Stack organization is actually a preparation procedure of substances in LIFO management scheme. This management scheme makes this easy to stock and restore any objective. Unlike other software design languages wherever the developer desires to assign data and assemble garbage, Java has the capability of automatic trash gathering and memory distribution.
4. Distributiveness: Distributive calculating is basically the stage where two or more processors can work composed on a net. Java has great networking competence. Interacting on Java is too informal that writing networking program senses like receiving and sending between files.
5. Open-Source: Java app expansion is an open source policy which permits the companies to download the apparatuses and develop apps without recompensing any authorization fee or yearly contribution amount; this helps in dropping the inclusive cost of the expansion process.
6. Platform Independent: Java is presented with platform independency conduct. Java applications could be advanced and usage in any operating system or podium. This is why the companies mostly favor Java applications since it offers an ease of using those applications on any OS.
7. Stack Object Allocation: The leading feature of Java application development is that this uses the stack distribution system for data using and storing. Stack object distribution management system is really LIFO system whichever means Last-In First-Out. If an objective is stored in the previous, that object would be accessed first. This is very helpful for the companies as they could easily entrance the data stored.
8. Extremely secure: Java is well recognized for its security standards and safe programming. The companies could download any folder with non-trusted programs and then the application can use these non-trusted and non-secure codes in a safe and secure way. The apps mechanically devoid the stores corrupted or infected and usage them in application
9. Reusable Codes: The companies could reuse the codes in emerging any other application on Java stage. The codes usage in the software design of any application could be reused by the companies in making any other application.
10. Nature and Behavior: The applications made on Java stand are very vibrant in nature. The codes usage in the programming isdecided in precise object-concerned with units to be reserved in a distinct file for respectively object code.
11. User and designer friendly: The companies are attracted to the Java development as the applications are not only stress-free to progress but very informal to usage also. A very influential and elegant set of APIs makes this easier to progress apps and produce codes rapidly which protects the time in addition to money of the companies (MRBOOL, 2014).

## 3.3 SYSTEM DESIGN

System design consists of design activities that produce system specifications satisfying the functional requirements that were developed in the system analysis process.

System design specifies how the system will accomplish the objectives form system analysis. System design is the structural implementation of the system analysis.

**SYSTEM REQUIREMENTS**

**Functional Requirements**

* The system shall include an authorization process where users can login their names and password.
* The system shall ensure that data on the system is secured.
* The system shall have a database.
* The system should generate report based on the raw data inputed

**Non-functional Requirements**

* The system shall be available in real time.
* The system shall be user friendly.
* The system shall be robust.
* The system shall be portable.

The system shall have an interface

### 3.4.2 USER REQUIREMENTS

To access the market basket analysis system, the user would require;

* The system shall allow authentication and grant access to valid users
* The system shall allow large volume of data from valid users
* The user should be able to save data in the database
* The user should be able to request for report

## 3.5 DESIGN APPROACH

### 3.5.1 USE CASE DIAGRAM

A use case diagram is used to show the functionality provided by a system in terms of actors, their goals -represented as use cases, and any dependencies between themIt depicts every functionality that an end user can do on the system to be developed. In other words, It represents every case of use or action

### SEQUENCE DIAGRAM

 This is the sequence of steps in achieving a use case goal that shows the interactions between the users and the system's major components. Sequence diagrams are used to depict the interactions between system objects or entities in the order of occurrence.



### FLOW CHART DIAGRAM

See Appendix A

# 3.6 DATABASE

A database is a structured collection of data for one or more purposes, usually in digital form. The data are typically structured to model relevant aspects of reality, in a way that supports processes requiring this information. The term "database" refers both to the way its users view it, and to the logical and physical materialization of its data, content, in files, computer memory, and computer data storage. Typically, for a given database, there is a structural description (schema) of the type of facts held in that database. There are a number of different ways of organizing a schema known as Database Models (Data Model).This definition is very general, and is independent of the technology used. However, not every collection of data is a database; the term database implies that the data is managed to some level of quality (measured in terms of accuracy, availability, usability, and resilience) and this in turn often implies the use of a general-purpose Database management system (DBMS). A general-purpose DBMS is typically a complex software system that meets many usage requirements, and the databases that it maintains are often large and complex. MySQL database which was adopted in a relational database management system (RDBMS) supported on all platforms. Java includes native support for data programming with MySQL database. It can be used to write and debug code to be executed by SQL. It also includes a data designer that can be used to graphically create, view or edit database schemas. Queries can be created either visually or using codes.

## 3.6.1 **DATABASE DESIGN**

The term database design refers to the overall process of designing, not just the base data structures, but also the forms and queries used as part of the overall database application within the [database management system](http://en.wikipedia.org/wiki/Database_management_system) (DBM). It is the process of producing the [logical data model](http://en.wikipedia.org/wiki/Logical_data_model)  which contains all the needed logical and physical design choices, and physical storage parameters needed to generate a design in a [Data Definition Language](http://en.wikipedia.org/wiki/Data_Definition_Language). It can also be thought of as the logical design of the base data structures used to store the data. In the [relational model](http://en.wikipedia.org/wiki/Relational_model) these are the [tables](http://en.wikipedia.org/wiki/Database_table) and [views](http://en.wikipedia.org/wiki/Database_view) which entails the description of the various table, fields that the tables contains and their corresponding data types as well as the field length. In the MySQL structure, the term database refers to group of tables (which consists of rows and columns containing the data for the tables) and other objects such as indexes. A database can contain virtually unlimited number of tables.

### 3.6.2 TABLES

Tables are one of the most important components of a database. This is because MySQL uses tables to store data thus one creates are defined. These defined are the individual pieces of information (fields).One can decide how you want information to be saved in a specific column by arranging the properties below:

* Default: The present option state value for a column (default value).
* Rule: The valid or logical value for data entry.
* Constraints: rules for confirming (validation) data entry.
* Data type: The type of data that can be entered into the column.

CHAPTER FOUR

SYSTEM IMPLEMENTATION AND TESTING

4.0 Introduction

This chapter presents the system flow chart, choice of programming language and programming environment.

4.3 PROGRAM TESTING

The aim here is to ensure that this program meets its requirements. A user can perform analysis in any of group of products. At this stage of the work every eventuality has been subjected to the most vigorous examination as envisage on the specification.

4.4 SOFTWARE TESTING

 Testing the software follows a certain process as shown below that occurs when errors are uncovered. As a secondary benefit, testing demonstrates that the software functions appear to be working according to specification, that behavioral and performance requirements appear to have been met. In addition, data collected as testing is conducted, provide a good indication of software reliability and quality as a whole.

* 1. SYSTEM DESIGN DIAGRAM

Main Menu

Login

Exit

Select File

Barchart

4.6 Choice of programming language

The programming language chosen for the development of the system is Java. the language was chosen because; it is pure object-oriented, it is more type safe, you need not put much attention on problems as memory leak and ease-to-development, the rich class library makes many functions easy to be implemented.

* 1. Programming Environment

This system can function effectively in windows such as; window 7, window 8 and windows 10 etc. However, because of the structural concept of the language used, it is flexible for amendment in future when need arises

* + 1. Hardware Requirement
* Super video graphic array monitor
* At least 512 MB RAM
* At least 40 GB hard disk
* Keyboard
* Mouse
* Uninterruptible Power supply (UPS)

4.7.2 Software Requirement

* .JRE
	1. System Implementation

Implementation is the process of replacing the old system with the new system. There are four different ways of replacing the old system with the new system. The reasons for choosing one implementation type over another depend upon; how quickly must of the changeover happen? How important is it to prevent data loss? What will the cost of the changeover be?

Direct changeover: In this system the old system is no longer available and everything must run on the new system. Problems with the new system can cause major problems for the business, only suitable for non-critical systems.

Phased implementation: Takes longer to complete the implementation but the risks to the business are less than for direct changeover. The new system can be split into separate working parts e.g. sales, marketing, payroll etc. part of the old system is replaced with the new one until the replaced part is working properly. Continue the process until the entire old system has been replaced by the new system.

Parallel Running: Highly fault tolerant, new system and the old system are used with extra staffs recruited to run the new system but it is very expensive. Both systems continue to run until the new system is working properly then the old one is discarded.

Pilot Running: If the business has many different offices or sites then this is an option. One single site is chosen and the old system is replaced with the new system in the same way as direct changeover but only on one site, the rest of the business continue to use the old system. Once the new system is shown to work well in that one ‘pilot’ site then the new system can replace the old one in the rest of the company.

The system implementation method recommended and chosen by the system developer is the parallel running so as to prevent data loss.

4.2 SYSTEM SUB FUNCTION IMPLEMENTATION

ENTRY MODULE

The entire module is responsible for all input data requirements. It receives input data from the computer users and stores them adequately into file. The entire provisions are made very flexible and precise:

REPORT MODULE

The report module, tables all forms of report generation. It displays conditional and unconditional reports. The reports are made comprehensive and timely:

UPDATE MODULE

The update module is responsible for modifying stored data or record in the files. The records are searched for in the file and retrieved adequately and then, the update data are retrieved and necessary corrections are made automatically by the computer as adequate. However, it is responsible in keeping track of all the data and record that takes place.

EXIT MODULE

This is the module responsible for packing up or quitting the program entirely:

CHAPTER FIVE

**5.0 Introduction**

This chapter focuses on summary, conclusion and recommendations.

Here, the entire summary of the research from the problem stage to the implementation stage, the relevant conclusion and recommendations are discussed.

**5.1 Constraints of the Study**

The problems encountered during the course of carrying out this research project include:

1. **Time**: Time for the research project was too short coupled with researcher’s academic time table.

2. **Fund**:There was limited fund to take care of the research properly.

3. **Research Materials:** Lack of access to research materials on the topic in the school library and even public libraries were also major constraint in the cause of this project.

**5.2 Summary**

Implementation ofmarket basket analysis has been carried out using Java. The manual method of analyzing sales has been so stressful and time consuming. The new system would be very easy to use because of its accuracy and reliability. Analysis report can be promptly assessed easily.

**5.3 Conclusion**

In conclusion, market basket analysis system has advantages and disadvantages. I believe that in my opinion, that the advantages outweigh the disadvantages. With our new developed system supermakets will be able to easily make complex decisions within minutes. Its low cost means it can be easily implemented thereby eliminating the cost of hiring professionals to perform any analysis. It proves to be the solution for most supermarkets and guarantees result every time.

**5.4 Recommendations**

Having designed, tested and implemented the new system, the following must be put in place to fully achieve the objective of which the software is designed.

1. **Maintenance:** The system needs to be maintained. This implies that any fault detected should be reported to the programmer for correction at any point in time.

**Research:** More research should be conducted on the topic to assess it effectively.

**REFERENCES**

Jiawei Han and Micheline Kamber, “Data Mining: Concepts andTechniques”, 2 edition (4 Jun 2006),

Gordon S. Linoff and Michael J. Berry, “Data Mining Techniques:For Marketing, Sales, and Customer Relationship Management”,3rd Edition edition (1 April 2011)

Vipin Kumar and Mahesh Joshi, “Tutorial on High PerformanceData Mining ”, 1999

Rakesh Agrawal, Ramakrishnan Srikan, “Fast Algorithms forMining Association Rules”, Proc VLDB, 1994

(<http://www.cs.tau.ac.il/~fiat/dmsem03/Fast%20Algorithms%20for%20Mining%20Association%20Rules.ppt>)

U. Fayyad et al. 1996 *Advances in Knowledge Discovery and Data Mining* [ISBN 978-0-262-56097-9](https://en.wikipedia.org/wiki/Special%3ABookSources/9780262560979)

APPENDIX A (PROGRAM FLOWCHART)

Is login correct?

Load .xls file with records for analysis.

Is choice select file?

Display barchart representation of analysis report

Is choice barchart ?

Is choice Exit ?

APPENDIX B (SOURCE CODE)

SOURCE CODE

package mba;

import java.awt.\*;

import java.awt.event.\*;

import java.io.File;

import java.io.IOException;

import java.util.Vector;

import java.util.logging.Level;

import java.util.logging.Logger;

import javax.swing.\*;

import javax.swing.filechooser.FileFilter;

import javax.swing.table.DefaultTableModel;

import jxl.Cell;

import jxl.Sheet;

import jxl.Workbook;

import jxl.read.biff.BiffException;

import static mba.excelTojTable.headers;

import static mba.excelTojTable.model;

import static mba.excelTojTable.table;

import org.jfree.chart.ChartFactory;

import org.jfree.chart.ChartPanel;

import org.jfree.chart.JFreeChart;

import org.jfree.chart.plot.CategoryPlot;

import org.jfree.chart.plot.PlotOrientation;

import org.jfree.data.category.DefaultCategoryDataset;

import java.sql.Connection;

import java.sql.DriverManager;

import java.sql.SQLException;

import java.sql.Statement;

import java.sql.ResultSet;

import java.util.Random;

import javax.swing.JFrame;

import javax.swing.JOptionPane;

import javax.swing.JScrollPane;

import javax.swing.JTable;

import javax.swing.table.DefaultTableModel;

import javax.swing.table.TableRowSorter;

public class analysis extends javax.swing.JFrame {

 Connection con;

Statement stmt;

ResultSet rs;

int curRow = 0;

String username;

 //JTable table;

 JScrollPane scroll;

 String u1;

 String u2;

 String u3;

 String u4;

 String u5;

 String u6;

 int k1;

 int k2;

 int k3;

 int k4;

 int k5;

 int k6;

 //DefaultTableModel model = null;

Vector headers = new Vector();

 Vector data = new Vector();

 int tableWidth = 0;

 int tableHeight = 0;

 JFileChooser jChooser;

 /\*\*

 \* Creates new form analysis

 \*/

 public analysis() {

 initComponents();

 }

 jButton3 = new javax.swing.JButton();

 jButton4 = new javax.swing.JButton();

 jButton5 = new javax.swing.JButton();

 jScrollPane3 = new javax.swing.JScrollPane();

 otable3 = new javax.swing.JTable();

 jButton7 = new javax.swing.JButton();

 jButton6 = new javax.swing.JButton();

 jButton8 = new javax.swing.JButton();

 jButton9 = new javax.swing.JButton();

 jScrollPane4 = new javax.swing.JScrollPane();

 otable4 = new javax.swing.JTable();

 jLabel1 = new javax.swing.JLabel();

 setDefaultCloseOperation(javax.swing.WindowConstants.DISPOSE\_ON\_CLOSE);

 addWindowListener(new java.awt.event.WindowAdapter() {

 public void windowActivated(java.awt.event.WindowEvent evt) {

 formWindowActivated(evt);

 }

 });

 jButton2.setText("Select file");

 jButton2.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton2ActionPerformed(evt);

 }

 });

 otable.setModel(new javax.swing.table.DefaultTableModel(

 new Object [][] {

 {null, null, null, null},

 {null, null, null, null},

 {null, null, null, null},

 {null, null, null, null}

 },

 new String [] {

 "Title 1", "Title 2", "Title 3", "Title 4"

 }

 ));

 jScrollPane1.setViewportView(otable);

 otable2.setModel(new javax.swing.table.DefaultTableModel(

 new Object [][] {

 },

 new String [] {

 "Products", "No. of times pruchased"

 }

 ));

 jScrollPane2.setViewportView(otable2);

 jButton1.setText("jButton1");

 jButton1.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton1ActionPerformed(evt);

 }

 });

 mypanel.setBackground(new java.awt.Color(153, 153, 153));

 mypanel.setLayout(new java.awt.BorderLayout());

 jButton3.setText("jButton3");

 jButton3.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton3ActionPerformed(evt);

 }

 });

 jButton4.setText("Bar chart");

 jButton4.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton4ActionPerformed(evt);

 }

 });

 jButton5.setText("chrt");

 jButton5.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton5ActionPerformed(evt);

 }

 });

 otable3.setModel(new javax.swing.table.DefaultTableModel(

 new Object [][] {

 },

 new String [] {

 "Product", "tyms"

 }

 ));

 jScrollPane3.setViewportView(otable3);

 jButton7.setText("jButton7");

 jButton7.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton7ActionPerformed(evt);

 }

 });

 javax.swing.GroupLayout jPanel1Layout = new javax.swing.GroupLayout(jPanel1);

 jPanel1.setLayout(jPanel1Layout);

 jPanel1Layout.setHorizontalGroup(

 jPanel1Layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addContainerGap()

 .addGroup(jPanel1Layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addGroup(jPanel1Layout.createParallelGroup(javax.swing.GroupLayout.Alignment.TRAILING)

 .addComponent(jButton4)

 .addComponent(jButton3))

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.UNRELATED)

 .addComponent(jScrollPane3, javax.swing.GroupLayout.PREFERRED\_SIZE, 44, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGap(0, 0, Short.MAX\_VALUE))

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addComponent(jButton5)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED, javax.swing.GroupLayout.DEFAULT\_SIZE, Short.MAX\_VALUE)

 .addComponent(jButton7)))

 .addContainerGap())

 );

 jPanel1Layout.setVerticalGroup(

 jPanel1Layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addContainerGap()

 .addComponent(jButton3)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.UNRELATED)

 .addGroup(jPanel1Layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addComponent(jButton4)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED)

 .addComponent(jButton5))

 .addGroup(jPanel1Layout.createSequentialGroup()

 .addComponent(jScrollPane3, javax.swing.GroupLayout.PREFERRED\_SIZE, 32, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.UNRELATED)

 .addComponent(jButton7)))

 .addContainerGap(javax.swing.GroupLayout.DEFAULT\_SIZE, Short.MAX\_VALUE))

 );

 jButton6.setText("Bar chart");

 jButton6.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton6ActionPerformed(evt);

 }

 });

 jButton8.setText("jButton8");

 jButton8.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton8ActionPerformed(evt);

 }

 });

 jButton9.setText("jButton9");

 jButton9.addActionListener(new java.awt.event.ActionListener() {

 public void actionPerformed(java.awt.event.ActionEvent evt) {

 jButton9ActionPerformed(evt);

 }

 });

 otable4.setModel(new javax.swing.table.DefaultTableModel(

 new Object [][] {

 },

 new String [] {

 "product", "tyms"

 }

 ));

 jScrollPane4.setViewportView(otable4);

 jLabel1.setFont(new java.awt.Font("Tahoma", 1, 12)); // NOI18N

 jLabel1.setForeground(new java.awt.Color(51, 153, 255));

 jLabel1.setText("Database of previous loaded data");

 javax.swing.GroupLayout layout = new javax.swing.GroupLayout(getContentPane());

 getContentPane().setLayout(layout);

 layout.setHorizontalGroup(

 layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(layout.createSequentialGroup()

 .addContainerGap()

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addComponent(mypanel, javax.swing.GroupLayout.PREFERRED\_SIZE, 727, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING, false)

 .addComponent(jScrollPane2, javax.swing.GroupLayout.PREFERRED\_SIZE, 727, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGroup(javax.swing.GroupLayout.Alignment.TRAILING, layout.createSequentialGroup()

 .addGap(218, 218, 218)

 .addComponent(jButton8, javax.swing.GroupLayout.PREFERRED\_SIZE, 0, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED, javax.swing.GroupLayout.DEFAULT\_SIZE, Short.MAX\_VALUE)

 .addComponent(jButton2)

 .addGap(53, 53, 53)

 .addComponent(jButton6)

 .addGap(134, 134, 134)

 .addComponent(jButton9, javax.swing.GroupLayout.PREFERRED\_SIZE, 0, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGap(65, 65, 65))))

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(layout.createSequentialGroup()

 .addGap(18, 18, 18)

 .addComponent(jScrollPane4, javax.swing.GroupLayout.PREFERRED\_SIZE, javax.swing.GroupLayout.DEFAULT\_SIZE, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGap(10, 10, 10)

 .addComponent(jPanel1, javax.swing.GroupLayout.PREFERRED\_SIZE, 10, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGap(484, 484, 484)

 .addComponent(jScrollPane1, javax.swing.GroupLayout.PREFERRED\_SIZE, 2, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addGap(475, 475, 475)

 .addComponent(lm))

 .addGroup(layout.createSequentialGroup()

 .addGap(133, 133, 133)

 .addComponent(jLabel1, javax.swing.GroupLayout.PREFERRED\_SIZE, 254, javax.swing.GroupLayout.PREFERRED\_SIZE)))

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED)

 .addComponent(jButton1, javax.swing.GroupLayout.PREFERRED\_SIZE, 0, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addContainerGap(javax.swing.GroupLayout.DEFAULT\_SIZE, Short.MAX\_VALUE))

 );

 layout.setVerticalGroup(

 layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(layout.createSequentialGroup()

 .addContainerGap()

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(layout.createSequentialGroup()

 .addComponent(jScrollPane2, javax.swing.GroupLayout.PREFERRED\_SIZE, 195, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED, 10, Short.MAX\_VALUE)

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(javax.swing.GroupLayout.Alignment.TRAILING, layout.createParallelGroup(javax.swing.GroupLayout.Alignment.BASELINE)

 .addComponent(jButton2)

 .addComponent(jButton8, javax.swing.GroupLayout.PREFERRED\_SIZE, 6, javax.swing.GroupLayout.PREFERRED\_SIZE))

 .addGroup(javax.swing.GroupLayout.Alignment.TRAILING, layout.createParallelGroup(javax.swing.GroupLayout.Alignment.BASELINE)

 .addComponent(jButton6)

 .addComponent(jButton9, javax.swing.GroupLayout.PREFERRED\_SIZE, 6, javax.swing.GroupLayout.PREFERRED\_SIZE))

 .addComponent(jPanel1, javax.swing.GroupLayout.PREFERRED\_SIZE, 0, javax.swing.GroupLayout.PREFERRED\_SIZE))

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.RELATED)

 .addComponent(mypanel, javax.swing.GroupLayout.PREFERRED\_SIZE, 379, javax.swing.GroupLayout.PREFERRED\_SIZE))

 .addGroup(layout.createSequentialGroup()

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addGroup(layout.createSequentialGroup()

 .addGap(603, 603, 603)

 .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)

 .addComponent(jScrollPane1, javax.swing.GroupLayout.Alignment.TRAILING, javax.swing.GroupLayout.PREFERRED\_SIZE, 3, javax.swing.GroupLayout.PREFERRED\_SIZE)

 .addComponent(jButton1, javax.swing.GroupLayout.Alignment.TRAILING, javax.swing.GroupLayout.PREFERRED\_SIZE, 3, javax.swing.GroupLayout.PREFERRED\_SIZE)))

 .addGroup(layout.createSequentialGroup()

 .addComponent(jLabel1)

 .addPreferredGap(javax.swing.LayoutStyle.ComponentPlacement.UNRELATED)

 .addComponent(jScrollPane4, javax.swing.GroupLayout.PREFERRED\_SIZE, 568, javax.swing.GroupLayout.PREFERRED\_SIZE)))

 .addGap(0, 0, Short.MAX\_VALUE)))

 .addGap(11, 11, 11)

 .addComponent(lm))

 );

 setSize(new java.awt.Dimension(1295, 674));

 setLocationRelativeTo(null);

 }// </editor-fold>

 private void jButton2ActionPerformed(java.awt.event.ActionEvent evt) {

 JFileChooser chooser = new JFileChooser();

DefaultTableModel model = (DefaultTableModel)otable.getModel();

chooser.setFileSelectionMode(JFileChooser.FILES\_ONLY);

chooser.setFileFilter(new FileFilter() {

 @Override

 public String getDescription() {

 return "All supported image formats";

 }

 @Override

 public boolean accept(File f) {

 if (f.isDirectory()) {

 return true;

 } else {

 return f.getName().toLowerCase().endsWith(".xls");

 }

 }

});

int res = chooser.showOpenDialog(analysis.this);

if (res == JFileChooser.APPROVE\_OPTION) {

 File file = chooser.getSelectedFile();

 fillData(file);

 model = new DefaultTableModel(data,

 headers);

 tableWidth = model.getColumnCount()

 \* 150;

 tableHeight = model.getRowCount()

 \* 25;

 otable.setPreferredSize(new Dimension(

 tableWidth, tableHeight));

 otable.setModel(model);

 uki();

 }

else {

 JOptionPane.showMessageDialog(analysis.this, "Dialog cancelled by the user");

}

 }

 private void jButton1ActionPerformed(java.awt.event.ActionEvent evt) {

 // TODO add your handling code here:

 DefaultTableModel model = (DefaultTableModel)otable.getModel();

 DefaultTableModel model2 = (DefaultTableModel)otable2.getModel();

 for (int i = model.getRowCount() - 1; i >= 0; i--){

 String lk = String.valueOf(model.getValueAt(i, 0));

 System.out.println(lk);

 // table2

 for (int p = model2.getRowCount() - 1; p >= 0; p--){

 String lk1 = String.valueOf(model2.getValueAt(p, 0));

 if (lk .equals(lk1)){

 String lk2 = String.valueOf(model2.getValueAt(p, 1));

 int uk = Integer.parseInt(lk2);

 uk = uk + 1;

 String kk = String.valueOf(uk);

 otable2.setValueAt(kk, p, 1);

 //p = -1;

 lm.setText("yes");

 }

 }

 if (lm.getText().equals("yes")){

 System.out.println("ok");

 lm.setText("no");

 }

 else{

 model2.addRow(new Object[] { lk, "1"});

 }

 }

 }

 private void jButton4ActionPerformed(java.awt.event.ActionEvent evt) {

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "Marks", u1);

 dataset.setValue(k2, "Marks", u2);

 dataset.setValue(k3, "Marks", u3);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Student score", "Student Name", "Marks", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void jButton3ActionPerformed(java.awt.event.ActionEvent evt) {

 DefaultTableModel model = (DefaultTableModel)otable2.getModel();

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 for (int i = model1.getRowCount() - 1; i >= 0; i--){

 model1.removeRow(0);

 }

 int k = 0;

 int f = otable2.getRowCount();

 f = f - 1;

 for (int i = f; i>-1; i--){

 Object cell = otable2.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable2.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 model1.addRow(new Object[] { valu, valu1});

 }

 }

 private void jButton5ActionPerformed(java.awt.event.ActionEvent evt) {

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>-1; i--){

 k = k + 1;

 if (k == 7){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 if (k == 4){

 u4 = valu;

 k4 = Integer.parseInt(valu1);

 }

 if (k == 5){

 u5 = valu;

 k5 = Integer.parseInt(valu1);

 }

 if (k == 6){

 u6 = valu;

 k6 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 dataset.setValue(k4, "No. of times purchased", u4);

 dataset.setValue(k5, "No. of times purchased", u5);

 dataset.setValue(k6, "No. of times purchased", u6);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void jButton6ActionPerformed(java.awt.event.ActionEvent evt) {

 // TODO add your handling code here:

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 if (f > 5){

 six();

 }

 if (f == 5){

 five();

 }

 if (f == 4){

 four();

 }

 if (f == 3){

 three();

 }

 if (f == 2){

 two();

 }

 if (f == 1){

 one();

 }

 }

 private void six(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 6){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 if (k == 4){

 u4 = valu;

 k4 = Integer.parseInt(valu1);

 }

 if (k == 5){

 u5 = valu;

 k5 = Integer.parseInt(valu1);

 }

 if (k == 6){

 u6 = valu;

 k6 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 dataset.setValue(k4, "No. of times purchased", u4);

 dataset.setValue(k5, "No. of times purchased", u5);

 dataset.setValue(k6, "No. of times purchased", u6);

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void five(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 5){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 if (k == 4){

 u4 = valu;

 k4 = Integer.parseInt(valu1);

 }

 if (k == 5){

 u5 = valu;

 k5 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 dataset.setValue(k4, "No. of times purchased", u4);

 dataset.setValue(k5, "No. of times purchased", u5);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void four(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 4){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 if (k == 4){

 u4 = valu;

 k4 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 dataset.setValue(k4, "No. of times purchased", u4);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void three(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 3){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void two(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 2){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void one(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 6){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 //System.out.println(valu);

 //System.out.println(valu1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void jButton7ActionPerformed(java.awt.event.ActionEvent evt) {

 // TODO add your handling code here:

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>0; i--){

 k = k + 1;

 if (k == 6){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 //dataset.setValue(k2, "No. of times purchased", u2);

 //dataset.setValue(k3, "No. of times purchased", u3);

 //dataset.setValue(k4, "No. of times purchased", u4);

 //dataset.setValue(k5, "No. of times purchased", u5);

 //dataset.setValue(k6, "No. of times purchased", u6);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

 }

 private void jButton8ActionPerformed(java.awt.event.ActionEvent evt) {

 // TODO add your handling code here:

 String connectionURL = "jdbc:derby:product/db";

 // String connectionURL = "jdbc:derby:onlinedb/exams";

 String createString = "CREATE TABLE login (unem varchar(100), pwod varchar(100))";

 try {

 con = DriverManager.getConnection(connectionURL);

 if (con != null){

 //con = DriverManager.getConnection(connectionURL);

 Statement stmt = con.createStatement();

 stmt.executeUpdate(createString);

 System.out.println("Connected to database- exams");

 }

 } catch (SQLException ex) {

 System.out.println(ex);

 }

 }

 private void jButton9ActionPerformed(java.awt.event.ActionEvent evt) {

 // TODO add your handling code here:

 String connectionURL = "jdbc:derby:product/db";

 try {

 con = DriverManager.getConnection(connectionURL);

 stmt = con.createStatement( ResultSet.TYPE\_SCROLL\_INSENSITIVE, ResultSet.CONCUR\_UPDATABLE );

 String SQL = "SELECT unem, pwod FROM login";

 rs = stmt.executeQuery( SQL );

 String u = "analysis";

 String u1 = "12345";

 rs.moveToInsertRow( );

 rs.updateString("unem", u);

 rs.updateString("pwod", u1);

 rs.insertRow( );

 stmt.close( );

 rs.close( );

 JOptionPane.showMessageDialog(analysis.this, "Registration was successful !!!");

 }

 catch (SQLException err) {

 System.out.println(err.getMessage() );

 }

 }

 private void formWindowActivated(java.awt.event.WindowEvent evt) {

 // TODO add your handling code here:

 String connectionURL = "jdbc:derby:product/db";

 try {

 //stmt.close();

 //rs.close();

 //String evid = st.getText();

 Connection con = DriverManager.getConnection(connectionURL);

 stmt = con.createStatement( ResultSet.TYPE\_SCROLL\_INSENSITIVE, ResultSet.CONCUR\_UPDATABLE );

 String SQL = "SELECT prod, times FROM mb";

 rs = stmt.executeQuery( SQL );

 DefaultTableModel model = (DefaultTableModel)otable4.getModel();

 for (int i = model.getRowCount() - 1; i >= 0; i--){

 model.removeRow(0);

 }

 while ( rs.next( ) ) {

 String ID = rs.getString("prod");

 String NAME = rs.getString("times");

 model.addRow(new Object[] { ID, NAME});

 }

 }

 catch (SQLException err) {

 JOptionPane.showMessageDialog(analysis.this, err.getMessage());

 }

 }

private void uki (){

 DefaultTableModel model = (DefaultTableModel)otable.getModel();

 DefaultTableModel model2 = (DefaultTableModel)otable2.getModel();

 for (int i = model.getRowCount() - 1; i >= 0; i--){

 String lk = String.valueOf(model.getValueAt(i, 0));

 System.out.println(lk);

 /////////////////// table2

 for (int p = model2.getRowCount() - 1; p >= 0; p--){

 // System.out.println(p);

 String lk1 = String.valueOf(model2.getValueAt(p, 0));

 //String lk2 = String.valueOf(model2.getValueAt(p, 1));

 if (lk .equals(lk1)){

 String lk2 = String.valueOf(model2.getValueAt(p, 1));

 int uk = Integer.parseInt(lk2);

 uk = uk + 1;

 String kk = String.valueOf(uk);

 otable2.setValueAt(kk, p, 1);

 //p = -1;

 lm.setText("yes");

 }

 //if (lm.getText().equals("no")){

 // model2.addRow(new Object[] { lk, "1"});

 // lm.setText("yes");

 // }

 }

 if (lm.getText().equals("yes")){

 System.out.println("ok");

 lm.setText("no");

 }

 else{

 model2.addRow(new Object[] { lk, "1"});

 }

 }

 //////

 inem();

}

private void inem (){

 DefaultTableModel model = (DefaultTableModel)otable2.getModel();

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 for (int i = model1.getRowCount() - 1; i >= 0; i--){

 model1.removeRow(0);

 }

 int k = 0;

 int f = otable2.getRowCount();

 f = f - 1;

 for (int i = f; i>-1; i--){

 Object cell = otable2.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable2.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 model1.addRow(new Object[] { valu, valu1});

 //database

 String connectionURL = "jdbc:derby:product/db";

 try {

 con = DriverManager.getConnection(connectionURL);

 stmt = con.createStatement( ResultSet.TYPE\_SCROLL\_INSENSITIVE, ResultSet.CONCUR\_UPDATABLE );

 String SQL = "SELECT prod, times FROM mb";

 rs = stmt.executeQuery( SQL );

 rs.moveToInsertRow( );

 rs.updateString("prod", valu);

 rs.updateString("times", valu1);

 rs.insertRow( );

 stmt.close( );

 rs.close( );

 }

 catch (SQLException err) {

 System.out.println(err.getMessage() );

 }

 }

 // inem2();

}

private void inem2(){

 DefaultTableModel model1 = (DefaultTableModel)otable3.getModel();

 int k = 0;

 int f = otable3.getRowCount();

 f = f - 1;

 String v;

 for (int i = f; i>-1; i--){

 k = k + 1;

 if (k == 7){

 i = 0;

 }

 Object cell = otable3.getValueAt(i, 0);

 String valu = String.valueOf(cell);

 Object cell1 = otable3.getValueAt(i, 1);

 //model.removeRow(i);

 String valu1 = String.valueOf(cell1);

 if (k == 1){

 u1 = valu;

 k1 = Integer.parseInt(valu1);

 }

 if (k == 2){

 u2 = valu;

 k2 = Integer.parseInt(valu1);

 }

 if (k == 3){

 u3 = valu;

 k3 = Integer.parseInt(valu1);

 }

 if (k == 4){

 u4 = valu;

 k4 = Integer.parseInt(valu1);

 }

 if (k == 5){

 u5 = valu;

 k5 = Integer.parseInt(valu1);

 }

 if (k == 6){

 u6 = valu;

 k6 = Integer.parseInt(valu1);

 }

 model1.removeRow(i);

 }

 DefaultCategoryDataset dataset = new DefaultCategoryDataset();

 dataset.setValue(k1, "No. of times purchased", u1);

 dataset.setValue(k2, "No. of times purchased", u2);

 dataset.setValue(k3, "No. of times purchased", u3);

 dataset.setValue(k4, "No. of times purchased", u4);

 dataset.setValue(k5, "No. of times purchased", u5);

 dataset.setValue(k6, "No. of times purchased", u6);

 //dataset.setValue(95, "Marks", "Student4");

 JFreeChart chart = ChartFactory.createBarChart("Market Analysis Using Bar Chart", "Products", "No. of times purchased", dataset, PlotOrientation.VERTICAL, false, true, false);

 CategoryPlot p = chart.getCategoryPlot();

 p.setRangeGridlinePaint(Color.BLACK);

 ChartPanel chartPanel = new ChartPanel( chart);

 mypanel.removeAll();

 mypanel.add(chartPanel, BorderLayout.CENTER);

 mypanel.validate();

}

 void fillData(File file) {

 Workbook workbook = null;

 try {

 try {

 workbook = Workbook.getWorkbook(file);

 } catch (IOException ex) {

 Logger.getLogger(

 excelTojTable.class.

 getName()).log(Level.SEVERE,

 null, ex);

 }

 Sheet sheet = workbook.getSheet(0);

 headers.clear();

 for (int i = 0; i < sheet.getColumns(); i++) {

 Cell cell1 = sheet.getCell(i, 0);

 headers.add(cell1.getContents());

}

 data.clear();

 for (int j = 1; j < sheet.getRows(); j++) {

 Vector d = new Vector();

 for (int i = 0; i < sheet.getColumns(); i++) {

 Cell cell = sheet.getCell(i, j);

 d.add(cell.getContents());

 }

 d.add("\n");

 data.add(d);

 }

 } catch (BiffException e) {

 e.printStackTrace();

 }

 }

 public static void main(String args[]) {

 try {

 for (javax.swing.UIManager.LookAndFeelInfo info : javax.swing.UIManager.getInstalledLookAndFeels()) {

 if ("Nimbus".equals(info.getName())) {

 javax.swing.UIManager.setLookAndFeel(info.getClassName());

 break;

 }

 }

 } catch (ClassNotFoundException ex) {

 java.util.logging.Logger.getLogger(analysis.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);

 } catch (InstantiationException ex) {

 java.util.logging.Logger.getLogger(analysis.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);

 } catch (IllegalAccessException ex) {

 java.util.logging.Logger.getLogger(analysis.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);

 } catch (javax.swing.UnsupportedLookAndFeelException ex) {

 java.util.logging.Logger.getLogger(analysis.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);

 }

 //</editor-fold>

 /\* Create and display the form \*/

 java.awt.EventQueue.invokeLater(new Runnable() {

 public void run() {

 new analysis().setVisible(true);

 }

 });

 }

 // Variables declaration - do not modify

 private javax.swing.JButton jButton1;

 private javax.swing.JButton jButton2;

 private javax.swing.JButton jButton3;

 private javax.swing.JButton jButton4;

 private javax.swing.JButton jButton5;

 private javax.swing.JButton jButton6;

 private javax.swing.JButton jButton7;

 private javax.swing.JButton jButton8;

 private javax.swing.JButton jButton9;

 private javax.swing.JLabel jLabel1;

 private javax.swing.JPanel jPanel1;

 private javax.swing.JScrollPane jScrollPane1;

 private javax.swing.JScrollPane jScrollPane2;

 private javax.swing.JScrollPane jScrollPane3;

 private javax.swing.JScrollPane jScrollPane4;

 private javax.swing.JLabel lm;

 private javax.swing.JPanel mypanel;

 private javax.swing.JTable otable;

 private javax.swing.JTable otable2;

 private javax.swing.JTable otable3;

 private javax.swing.JTable otable4;

 // End of variables declaration

}