UNIVERSITY OF NAIROBI

SCHOOL OF COMPUTING & INFORMATICS

Model to Determine Bank Teller Requirements and Predict Transactions Case for: Banking Industry

By

Chege Kennedy Ruewel Muchai

P58/61543/2010

Supervisor: Lawrence Muchemi

OCTOBER 2012

Report submitted in partial fulfillment of the requirement of the Masters of Science in Computer Science at the University of Nairobi

DECLARATION

This project as presented in this report is my original work and it has not been presented to any institution of higher learning for the purpose of an academic evaluation whatsoever.

Student Name: Kennedy Ruewel Muchai Chege

Student Number: P58/61543/2010

Signature: .... Date: ...22/October 2012...

This project has been submitted as partial fulfillment of the requirements of the Masters of Science degree in Computer Science of the University of Nairobi with my approval as the University Supervisor

Lawrence Muchemi G

School of Computing and Informatics

University of Nairobi

Signature: .... Date: ...22/10/2012...

DEDICATION

I wish to dedicate this work to my lovely parents Mr. and Mrs. David Chege Muchai, who offered me generous support throughout my project work. I wish also to dedicate this project to my lovely wife Mrs. Miriam Muchai for her support, encouragement, synergy, and great care not forgetting the late nights she had to endure during my masters' course. I wish also to dedicate this work to my lovely sisters who were there for me during my project work.

ACKNOWLEDGEMENT

I would like to take this opportunity to greatly thank my God for His grace, care, energy, good health and better life offered to me throughout my masters' course. I wish to appreciate the good conducive environment offered by the University of Nairobi (School of Computing & Informatics) and sincerely appreciate the overwhelming support of my project supervisor, Mr. Lawrence Muchemi for his valuable time and efforts extended towards completing this project. I would wish to appreciate my employer for the financial empowerment towards realizing the cost of this undertaking.

ABSTRACT

Customer Satisfaction is paramount to any business or industry; this is particularly felt in the Banking industry especially in Kenya where there is fierce competition among the players. The banks have conventionally been associated with queues that anytime one has to visit the banking hall, the thought of long wait in the queue deter them away. This has had a major impact on Customer Service.

Simulation was applied to model the current scenario and estimate performance metrics. Some scenarios were considered to find out how the existing system operated.

Resource utility and customer waiting time were used to evaluate the performance of the queuing system. A mathematical model based on mathematical theory of queues, Little's result, theorem, lemma, law or formula, expressed algebraically as: $L = \lambda W$ was developed in-line with the bank's standard on customer waiting times.

The teller staffing model was tested using ARENA simulation software and the result was a reduction of 60% in customer waiting time. The model provided a dynamic schedule solution that allocated tellers based on the customer arrivals/demand. DTREG was used to model the transaction data and build the decision trees that were used to predict workload/transactions for any given working day of the week.

Performance of queuing systems is dictated by the Input Source, Service System and the Queuing Discipline. The Input Source is categorized as Static or Dynamic. Most queuing systems are founded on Static arrival process where the probability of arrivals is described as the number of customers arriving per unit of time. This study however proposed a Dynamic approach where the Service Rate was determined by both the service facility (Tellers) as well as the arrivals (Customers). The service facility adjusted its capacity to match the changes in demand intensity by varying the staffing levels at different timings of the day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CBS – Core Banking System

NSE – Nairobi Stock Exchange

GoK – Government of Kenya

ATMs – Automated Teller Machines

LoS – Length of Stay

FTE – Full Time Equivalent

QoS – Quality of Service

FCFS – First Come, First Served

LCFS – Last Come, First Served

SIRO – Service in Random Order

VIP – Very Important Person

VAT – Value Added Tax

TAT – Turn-Around Time

IPO – Initial Public Offering

WAN – Wide Area Network

L – Length of Queue

λ – Average Arrival Rate

W – Average waiting time

VB – Visual Basic 6.0

Chapter 1 : INTRODUCTION

1.1 Background

The banking industry is highly competitive, with banks not only competing among each other; but also with non-banks and other financial institutions [Kaynak and Kucukemiroglu, 1992; Hull, 2002]. Most bank product developments are easy to duplicate and when banks provide nearly identical services, they can only distinguish themselves on the basis of price and quality.

The key factors influencing customers' selection of a bank include: - the range of services, rates, fees and prices charged. It is apparent that superior service, alone, is not sufficient to satisfy customers [Abratt and Russell, 1999]. Prices are essential, if not more important than service and relationship quality. Furthermore, service excellence, meeting client needs, and providing innovative products are essential to succeed in the banking industry

In spite of the availability of Automated Teller Machines (ATMs), many customers still prefer to use human teller services, but long wait for service is perceived as a major source of customer dissatisfaction. [Cohen, Gan etl, 2006] The banks have attempted to provide quick service, whether in the bank's lobby or outside the banking hall facility by providing the demanded service while monitoring teller manning costs. Management's task is to have enough teller stations open to provide quick service while ensuring productive work time is not wasted by having idle tellers.

The business environment/climate in which organizations operate today is ever changing, and it's becoming more and more complex. Organizations both public and private, feel increasing pressures that force them to respond quickly to changing conditions and to be innovative in the way they operate. Such activities require organizations to be agile and make frequent and quick strategic, tactical and operational decisions some of which are very complex. Making such decisions may require considerable amounts of relevant data, information, and knowledge. Processing this in the framework of the needed decisions must be done quickly, frequently, in real-time, and usually requires some computerized support [Turban and Aronson, 1997]

Increasing productivity of Banking Operations has become a major issue in Bank management. Teller line services have been identified as the major area for productivity improvement. This fact has highlighted the need for establishing optimal staffing levels based on standards of customer service. [Pihl and Wambay, 1990] Customers demand higher standards of service and now have numerous choices from where to get served.

A common feature of many service industries ranging from telephone call centers to police stations and hospital emergency rooms is that, the demand for service often varies greatly by time of day.

1.2 Problem Statement

Queues are an everyday occurrence in most of our social lives for example while paying utility bills, supermarkets, banking halls etc. the most frustrating experience of a queue is the amount of time one has to spend while waiting to get served. This problem has been studied and solved in alternatives such as electronic ticketing where customers are issued with a numbered ticket on entering the banking hall, rest at a siting area provided until their number is called up. This however does not reduce the amount of time they'll spend waiting, the advantage of this model is that the customer won't be standing while waiting to be served. Another solution offered is to physically manage the customer traffic and call for additional resources to handle any demand and relive the resources when the demand level goes down. This theorem proposes a dynamic approach to managing these queues in terms of deploying the required resources to handle the demanded service

Setting staffing requirements is one in a hierarchy of decisions that must be made in the design and management of a service system [Turban and Aronson, 1997]. Customer satisfaction being a measurement of customer attitudes about products, services and brands is a growing concern in all industries more so service delivery settings. Customers expect to be served promptly when they arrive, and therefore there is a need for Optimal Staff Deployment on the customer facing operations. The Teller line especially should take into account varying service demand levels. From an internal survey conducted by the Bank in 2009, **long queues were highlighted to be the major source of customer dissatisfaction**. This translated to the long wait for service that Bank patrons had to endure.

This was also reflected by the customers' feedback and constant complaints that were received via the suggestion boxes, online social media networks such as Tweeter and Facebook.

1.3 Research Objectives

The idea is to have a Teller staffing model, based on work volumes and customer arrivals. This study aimed to achieve the following objectives:-

1. Develop a model that can be used to determine teller requirements
2. Predict transaction volumes for each day of a week.
3. Reduce the customer waiting time for service

1.4 Research Questions

This study sought to address waiting time for services and prediction of transactions and in so doing answer the following research questions:-

- What factors affect the performance of a queuing system?
- How can long waits for service be minimized?
- How can one predict customers' banking behavior?

1.5 Research Outcome

The solution that this research proposes was in essence a Decision Support Systems (DSS). In the early 1970s; Scott-Morton first articulated the major concepts of DSS. He defined DSS as "interactive computer-based systems which help decision makers utilize data and models so solve unstructured problems" [Gorry and Scott-Morton, 1971]. Another classic DSS definition provided by Keen and Scott-Morton: DSS systems couple the intellectual resources of individuals with the capabilities of computer to improve the quality of decisions. It is a computer-based support system for managerial decision makers who deal with semi structured problems [Keen and Scott-Morton (1978)]

The end-product was a computer-based model which made use of quantitative data that was used to measure the queue system performance and recommend an optimal number of personnel to be deployed at the teller system at any given time of the day as per the fluctuating customer demands. This henceforth assisted the decision maker/manager who is tasked with determining the manning requirements for the branch, to deploy the optimal number of staff. Collected transactional data was also be used to predict teller transaction volumes expected on various days of the week.

1.6 Significance of the Study

Tellers have huge impact on how customers feel about a financial institution. According to a survey carried out in 2010 in America that was entitled "Prime Performance 2010 Bank & Credit Union Satisfaction survey", that polled more than 6,000 customers in American banks, the message was clear that for most customers; tellers don't just represent the bank. They are the bank. Other than individual attitudes and behavior there are other factors that affect customer satisfaction for example, transaction speed and accuracy, wait time, and friendliness of staff

Tellers play a vital role in:

- Increasing productivity
- Improving service quality
- Maximizing sales and revenue possibilities

This research was significant to: -

- The student in the use of the knowledge acquired during the coursework to solve a real world problem
- The customers as long waits in banking halls were reduced which translated to banking being a delightful experience
- The Bank's Management as business growth was guaranteed as customer satisfaction was the key driver
- The Tellers as their productivity was increased since they handled the customers with ease despite their erratic arrivals

Based on this, there was the need to optimally deploy resources as; overstaffing increased operational costs while understaffing impacted negatively to the quality of customer service.

Why automate the process of knowing the Teller manning requirements

Quality support: - computers can improve the quality of the decision made as more data can be accessed, more alternatives can be evaluated, forecast can be improved, and risk analysis can be performed quickly. With computers, decision makers can perform complex simulations, check many possible scenarios and assess impact quickly and economically. Optimality is reached when (a mix of) objectives and performance measures (is) are satisfied. [Stegeman and Jansen-Vullers, 2005]

Agility support: - competition currently is based not just on price but also on quality, timeliness, customization of products and customer support

This research demonstrated the ideal Teller staffing required to cope with actual customer arrivals (demand for service) at any given time of the day.

The Banking services demand is more variable and often depends on

- The day of the week (Monday, Tuesday… Saturday)
- Month of the year (January, February… December)
- Time of day (peak and non-peak)
- Season e.g. public holidays, school reopening dates, IPOs, etc.

Service is delivered when and where it is needed. The goal of managers is to deploy as few employees as possible while maintaining high customer service standards. It is critical to recognize peak and non-peak periods to come up with the optimal number of Tellers to deploy. However identifying the challenges of demand levels was not straight forward. During off peak times the Tellers would stay idle and that yields to inefficient utilization of resources which has a direct undesirable impact on operational cost as the income is not commensurate to the staff remuneration. On the other hand understaffing the Teller line will result to poor customer service, overworked and unmotivated staff.

Unlike products that are stored in warehouses for future consumption, **a service is an intangible personal experience that cannot be transferred from one person to another**. Instead, a service is produced and consumed simultaneously. Whenever demand for a service falls short of the capacity to serve, the results are idle servers and facilities [Green, Kolesar and Whitt. (2007)].

The variability in demand is quite pronounced, and in fact, our culture and habits contribute to these fluctuations. For example, most of us eat our meals at the same hours, take our vacations in April and December, and pay school fees in January

Chapter 2 : LITERATURE REVIEW

2.1 Related Works

Personnel or staff scheduling problems have been studied for many years due to its importance on the overall performance of a system in terms of quality of service to the customer and cost to the organization. [Snell and Bohlander, 2007]

Donald Hammond and Sathi Mahesh (1995) using simulation and analysis proposed models to find cost effective bank teller management policies for providing high quality service levels at reasonable costs in a modern banking system.

Leeds (1992), while assessing quality of service to depend heavily on the quality of personnel, documented that approximately 40 percent of customers switched banks because of what they considered to be poor service. Leeds further argued that nearly three-quarters of the banking customers mentioned teller courtesy as a prime consideration in choosing a bank.

Fornell (1992), in his study of Swedish consumers, noted that although customer satisfaction and quality appear to be important for all firms, satisfaction is more important for loyalty in industries such as banks, insurance, mail order, and automobiles.

Gans (2003) developed a tool, Workforce Management (WFM) which is particularly suitable for day-to-day operations but does not answer the question of long-term planning

Ernst et al. (2004) used queuing models and simulation to obtain ideal staff requirements. Other authors mention this option as well, for example Mehrota (1997) and Grossman et al. (1999). Pichitlamken et al. (2003)

Chan (2003) identified Key Output Performance Variables (KOPV) and the Key Input Performance Variables (KIPV) which he used to design an effective workflow for call centers using simulation tools.

2.1.1 Bank Teller

The Oxford dictionary defines a Teller as a person employed to deal with customers' transactions. In some places, this employee is known as a cashier. Most teller jobs require cash handling experience. [Oxford dictionary]

Tellers are considered a "front line" in the banking business. This is because they are the first people that a customer sees at the bank and are also the people most likely to detect and stop fraudulent transactions in order to prevent losses at a bank (counterfeit currency and checks, identity theft, etc.). The position also requires tellers to be friendly and interact with the customers, providing them with information about customers' accounts and bank services.

The present words "tell" and "teller", are both based on an Old English word, "tellan," which in turn is based on a similar Germanic cognate. The original definition of tellan was "to reckon, calculate, count, consider or account." Over time, the word has evolved into "tell." The word teller dates back to around the end of the 15th century, with the traditional definition "person who counts." It eventually came to refer to a Bank Teller.

2.1.2 Human Resource Management (HRM)

HRM is the basis for all management activity but is not the basis for all business activity. A business may depend on having a unique product/service, or having necessary funding. The basis of management is always the same: getting the people of the business to make things happen in a productive way so that the business prospers and the people thrive. [Snell and Bohlander, 2007]

Managing resourceful humans requires a constant balancing between meeting the human aspirations of the people and meeting the strategic and financial needs of the business

By the end of the twentieth century most companies took towards the direction of ‘downsizing’ or reducing the number of people employed to create businesses that were lean, fit and flexible. Reducing the headcount became a fashionable criterion for success. Cost cutting achieved impressive short-term results, but it cannot be repeated year after year without impairing on the basic viability of the business.

There is now a move towards redressing that balance in search of equilibrium between the needs for financial viability and success in the market place on one hand and the need to maximize human capital on the other

Organizational effectiveness which is primarily defined in terms of meeting a service need as cost effective as possible and to the highest achievable standard of quality. [Snell and Bohlander, 2007] This defines the critical role that human resource play in any establishment

2.1.3 Technological Advancement

Advances in Information Technology (IT) have enabled organizations to take advantage of the information explosion. With computer networks, unlimited amounts of data can be stored, retrieved, and used in wide variety of ways. This has tremendously changed the way banks carry out most operations. In the 90’s for example, most teller operations were purely manual, this was the era of Passbook, where the customer balance was written in a piece of card that would have to be carried all time the customer visits the bank and would be manually updated on every transaction performed. This adversely affected the customer waiting times as simple operations could take extremely long time, consider a customer of a bank that has several branches (as is the model with banks) and wanted to transact on a different branch other than his/her domicile branch. This called for phone calls to be made so as to ascertain the balances and eventually serve the patron. Modern trends such as Online banking, mobile banking, etc. have revolutionized this via the use of sophisticated Online Banking systems that are accessible from the vast bank's Wide Area Networks (WAN) and in real time.

These technologies have significantly reduced the waiting times for service though it can still be minimized even further as proposed by this model

2.2 Queuing Theory

Queuing Theory is defined as a collection of mathematical models of various queuing systems. Used extensively to analyze production and service processes exhibiting random variability in market demand (arrival times) and service times [Adan and Resing, 2002].

Why do Queues Form?

Queues form as a result of the following:-

1. When the demand for a service facility exceeds the capacity of that facility i.e. the customer do not get served immediately upon arrival but must wait.
2. Some customers wait when the total number of customers requiring service exceeds the number of service facilities, some service facilities stand idle when the total number of service facilities exceeds the number of customers requiring service.

Waiting lines or queues are a common occurrence both in everyday life and in variety of business and industrial situations. **Most waiting line problems are centered about the question of finding the ideal level of services that a firm should provide**. For example

- Supermarkets must decide how many cash register check-out positions should be opened.
- Gasoline stations must decide how many pumps should be opened and how many attendants should be on duty.
- Manufacturing plants must determine the optimal number of mechanics to have on duty in each shift to repair machines that break down.
- Banks must decide how many teller windows to keep open to serve customers during various hours of the day.

Evolution of Queuing Theory

Queuing theory has its beginning in research work of a Danish engineer, A.K. Erlang. In 1990, Erlang experimented with fluctuating demand in telephone traffic. His work has since been extended to general problems and to business application of waiting lines.

Queuing Examples

Situation	Arriving customers	Service facility
Banking transactions	Bank patrons	Bank Tellers
Supermarket transactions	Shoppers	Cashiers
Scheduling of patients	Patients	Medical
Flow of computer programs through a computer system	Computer programs	Transmission lines

Figure 2.1: Queuing Examples

2.2.1 Components of a Basic Queuing Process

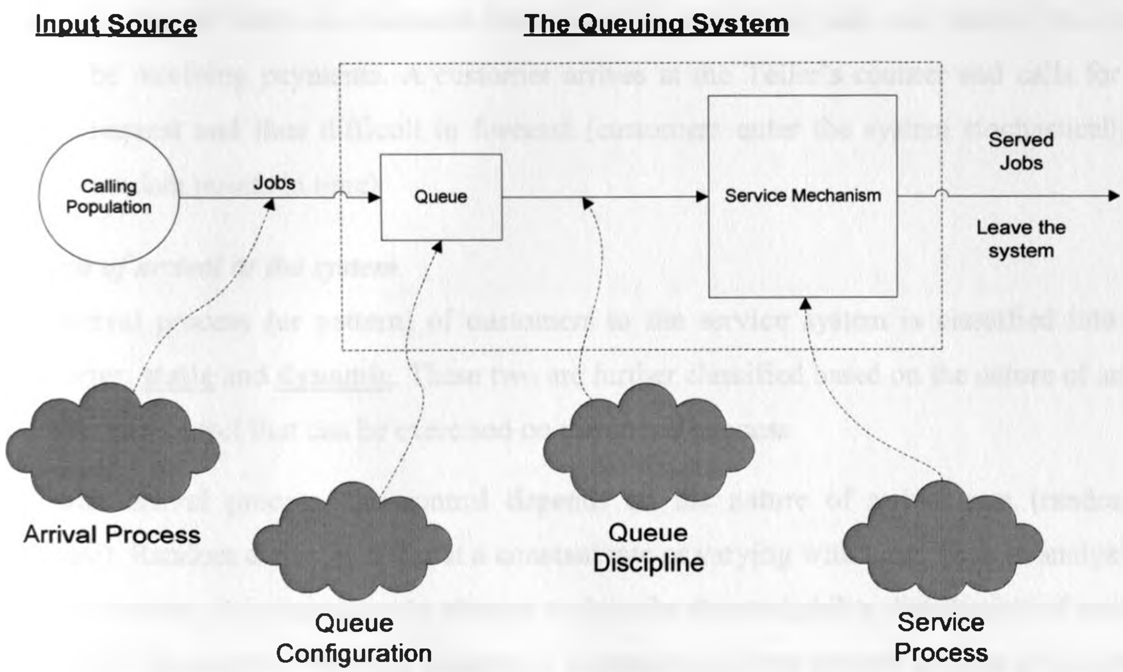


Figure 2.2: Components of a Basic Queue process

Input Source of Queue is characterized by:-

- Size of the calling population
- Pattern of arrivals at the system
- Behavior of the arrivals

Customers requiring service are generated at different times by an input source, commonly known as population. The rate at which customers arrive at the facility is determined by the arrival process

Size of the calling population

The size represents the total number of potential customers who will require service

1. According to source: - the source of the customers can be finite or infinite, in this case the population was finite as we focused on the bank's customers only.
2. According to numbers: - customers can arrive at a service facility according to some known schedule (for example one customer arrives every 5 minutes) or else they arrive randomly
3. According to time: - arrivals are considered random when they are independent of one another and their occurrence cannot be forecasted exactly, this was ideal for this research area as Bank patrons arrival cannot be clearly predicted and the demand for the different services was difficult to predict, for example one cannot tell that a Bank Teller will serve ten customer who would be depositing cash and twenty who would be receiving payments. A customer arrives at the Teller's counter and calls for any request and thus difficult to forecast (customers enter the system stochastically, at random points in time)

Pattern of arrival at the system

The arrival process (or pattern) of customers to the service system is classified into two categories: **static** and **dynamic**. These two are further classified based on the nature of arrival rate and the control that can be exercised on the arrival process

In ***static*** arrival process, the control depends on the nature of arrival rate (random or constant). Random arrival is either at a constant rate or varying with time. Thus to analyze the queuing system, it is necessary to attempt to describe the probability distribution of arrivals, also called inter-arrival time (i.e. number of customers arriving per unit of time at the service system)

The ***dynamic*** arrival system is controlled by both service facility and customer. The service facility adjusts its capacity to match changes in demand intensity, by varying the staffing levels at different timings of service, varying service charges at different timings, or allowing entry with appointments.

2.2.2 Service System

The service is provided by a service facility. In this instance it is a person (a Bank Teller). There are two aspects of a service system

a. The configuration of the service system and
b. The speed of the service

a. Configuration of the Service System

The customers' entry into the service system depends upon the queue conditions. If at the time of the customer's arrival, the server is idle, then the customer is served immediately. Otherwise the customer joins the queue which has several configurations i.e. how the service facilities exist. Service systems are classified by number of channels or servers (Tellers)

Single Server – Single Queue

This model involves one queue – one service station facility called single server models. The customer waits till the service point is ready to take him/her for servicing

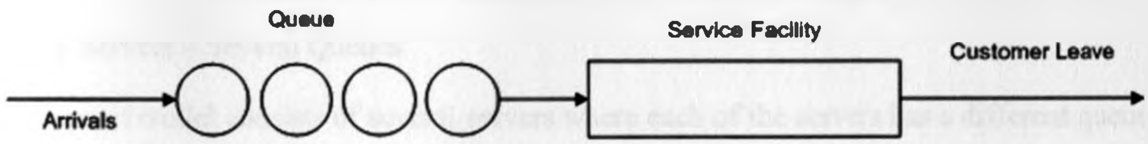


Figure 2.3: Single Server – Single Queue Model

Single Server – Several Queues

Here there are several queues and the customer may join any one of this but there is only one service channel

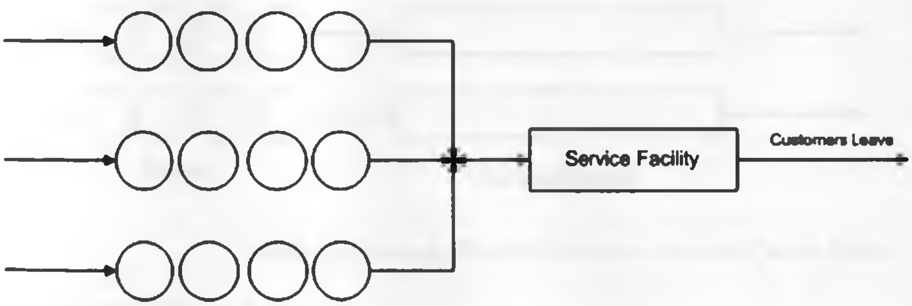


Figure 2.4: Single Server – Single Queue Model

Several (parallel) Servers – Single Queue

In this type of model there is more than one server and each server provides the same type of facility. The customer wait in a single queue until one of the service channels is ready to take them in for servicing. This was the queuing model that was used in this research as shown in Figure 2.5 below

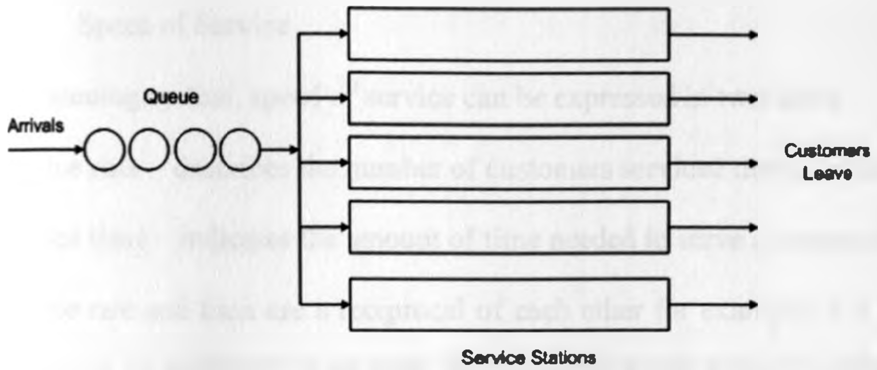


Figure 2.5: Several, Parallel Servers – Single Queue Model

Several Servers – Several Queues

This type of model consists of several servers where each of the servers has a different queue. A typical Kenyan example is the cash counters at Electricity House where customers can make payments of their electricity bills at any counter.

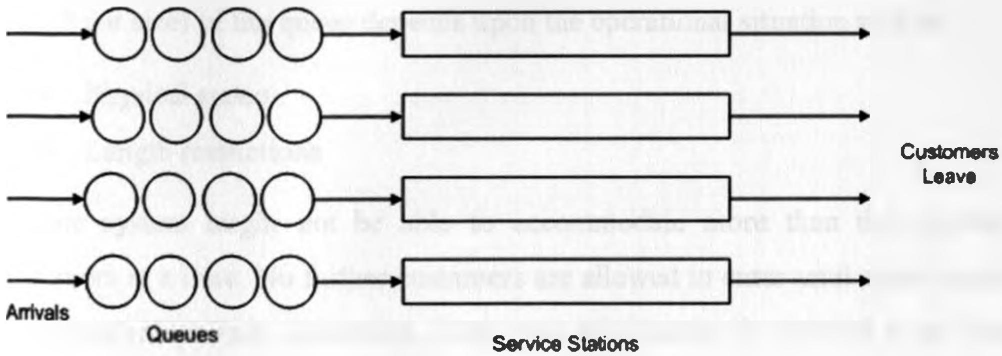


Figure 2.6: Several, Parallel Servers – Several Queue Model

Service Facilities in Series

In this model, a customer enters the first station and gets a portion of the service and then moves on to the next station, gets some service, continues like that and finally leaves the system, having received complete service. An example of this is a typical factory setting or an assembly plant where a vehicle is assembled in parts at several stages and exits as a complete unit

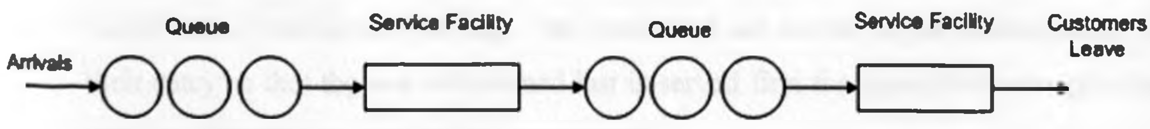


Figure 2.7: Multiple Servers in a Series

b. Speed of Service

In a queuing system, speed of service can be expressed in two ways

Service rate – describes the number of customers serviced during a particular time period

Service time – indicates the amount of time needed to serve a customer

Service rate and time are a reciprocal of each other for example, if a Teller can attend to an average of 20 customers in an hour, then the service rate would be expressed as, 20 customers / hour and service time would be equal to 3 minutes/customer [Adan and Resing. (2002)].

Queue Configuration

The queuing process refers to the number of queues, and their respective lengths. The number of queues depends on the layout of a service system thus there may be a single or multiple queues.

Length (or size) of the queue depends upon the operational situation such as

- Physical space
- Length restrictions

Service system might not be able to accommodate more than the required number of customers at a time. No further customers are allowed to enter until space becomes available to accommodate new customers. Such type of situation is referred to as finite or limited source queue for example cinema halls. On the other hand, a service queue can accommodate any number of customers at a time, referred to as infinite or unlimited source queue

2.2.3 Queue Discipline

In the queue structure, an important thing to know is the queue disciplines i.e. the order or manner in which customers from the queue are selected for service. This include

Static Queue Disciplines which are based on the individual customer's status in the queue system. A few of such disciplines are

i. First-Come, First-Served (FCFS) – customers are served in the order of their arrival
ii. Last-Come, First-Served (LCFS) - the customers are served in the reverse order of their entry so that the one who joined last is served first for example the people who enter an elevator last, are the first ones to exit
iii. Dynamic Queue Disciplines are based on the individual customer attributes in the queue. These are: -

a. Service in Random Order (SIRO) - under this rule, customers is selected for service at random, irrespective of their arrival in the service system. Every customer in the queue is equally likely to be selected. The time of arrival of the customer is therefore of no relevance
b. Priority Service – under this rule customers are grouped in priority classes on the basis of some attributes such as service time, urgency or to some identifiable characteristic. FCFS rule is used within each class to provide service.

2.3 Conceptual Model

Consider service system at many commercial Banks, where a customer, upon filling the relevant stationery queues and wait for the next available Teller, they are served and after that, exit the system. When there is no one in the queue, the customer is served immediately and exits the system as depicted in the figure below

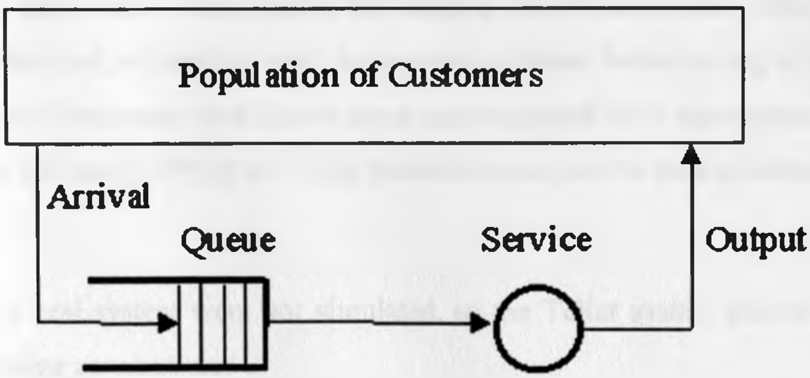


Figure 2.8: The Queuing System

Population: - Bank's customers

Arrivals: - random arrival rates varying with time as depicted in section 2.2.1

Queue: - several Parallel servers, single queue as illustrated in Fig 2.5 section 2.2.2

Queue discipline: - FCFS section 2.2.3

Service: - service times (durations for each transaction type) section 2.2.2b

Conceptual Model Objective

The Model intended to evaluate effectiveness of Bank Tellers and provide high quality service levels at reasonable costs. Bank's management usually wants to deploy adequate Tellers so that customers do not have to wait in line too long. An expected result is an optimum number of Tellers.

Inputs

Key inputs to the model were:-

- Service time statistics for the different transactions
- Number of transactions processed in a day
- Number of Service lines/facilities/Tellers
- Customer arrival rate

Content and Scope

Tellers did not have fixed lunch hours, but stepped out of the counter from time to time. However, a Teller had to finish serving the current customer before taking a break. Random customers entered the system at different times and requested for a transaction. Cash receipts (deposits) cater for nearly 75% of all Teller transactions as per the data collected.

Assumptions

All aspects of a real system were not simulated, so the Teller system had to be considered under the following assumptions

1. The working hours (Monday-Friday, 8.30 AM to 4 PM) with a 1 hour lunch break
2. Breaks – 30min
3. All Tellers were universal i.e. can perform any of the above mentioned function
4. All Tellers took the same time to process a transaction
5. All Tellers had equal capabilities
6. There was no system downtime
7. A customer was served by one Teller at a time
8. A customer cannot renege i.e. once the customer joins the queue they cannot leave without being served
9. Customers waited for service in a single queue
10. Customers were attended to on a first-come-first-served basis.

2.3.1 Conceptual Model Logic flow

Figure 2.9 below represents the logic flow as illustrated in the model description above

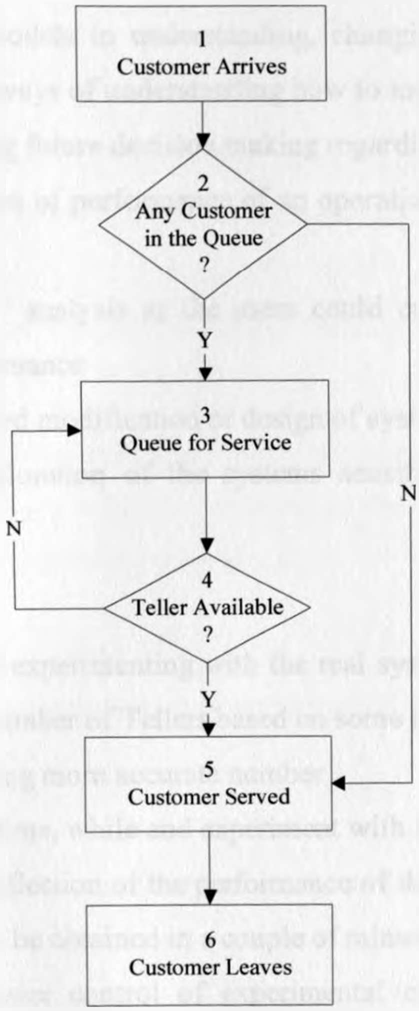


Figure 2.9: Logic flow for Parallel Servers, Single queue

2.3.2 Simulation

Simulation is the experimentation with a simplified imitation (on a computer) of an operations system as it progress through time, for the purpose of better understanding and/or improving that system [Robinson Stewart, 2004]. The main objective of the simulation was to show the running of the current system under a given set of parameters and to determine the optimal number of Tellers to deploy at any given time of day using performance metrics such as average length of the queue. We analyzed the distribution of the transactions on an hourly interval in order to establish the performance of the queuing system

Why Simulation?

1. To gain insights into when to make decisions
2. To make use of models in understanding, changing, managing and controlling reality. It involves ways of understanding how to improve the system
3. It enabled informing future decision making regarding the real system
4. It enabled prediction of performance of an operations system under a specific set of factors
5. To allow "what-if" analysis as the users could enter a scenario and the model predicted its performance
6. Simulation permitted modification or design of systems by trial and error
7. Allowed easy exploration of the systems sensitivity to changes in the input parameters

Advantages of Simulation

1. It is less costly, while experimenting with the real system, it can be more costly e.g. deploying a specific number of Tellers based on some previous experience as opposed to simulating and getting more accurate number
2. Simulation takes less time, while and experiment with a real system may take a longer duration before true reflection of the performance of the system can be obtained, with simulation, results can be obtained in a couple of minutes
3. Simulation allows easier control of experimental conditions, which is useful in comparing alternatives

Disadvantages of Simulation

1. Simulation software can be expensive as its software can be costly
2. Simulation requires expertise to project the real system

Modeling: - the area of interest must be simplified in order to be studied. A model had to take into account the relevant features of the real world.

Computer simulation are guided by **theory** as well as **experimental results**

Modeling and simulation aims to build a computer model that mimics the behavior of a complex real-life system so that one can better understand and optimize it.

In this research, using Modeling and Simulation we wanted to understand how Teller staffing levels impact on customer service, maximize use of resources during peak and off peak hours/seasons.

Discrete Event Simulation

For example, a Bank Teller serves patrons who arrived at random intervals and took up varying duration while being served. A customer's state at any time is either *waiting* or *being served* and the Teller sate at any point is either *serving* or *waiting*. By building a model of this Bank queue, we intended to simulate the effect of adding more Tellers during the busiest times.

By controlling the parameters of the modeled system, we intended to optimize on the number of Tellers in order to maintain customer satisfaction and minimize cost. Appropriate distribution model were used for the events simulated.

The Events

The Banking system is the real-life system. A system in general is a collection of entities which are logically related and which are of interest to particular application [Robinson Stewart, 2004].

The first step was to identify the basic events, whose occurrence would alter the status of the system. The Customer Arrival and departure events, are illustrated below

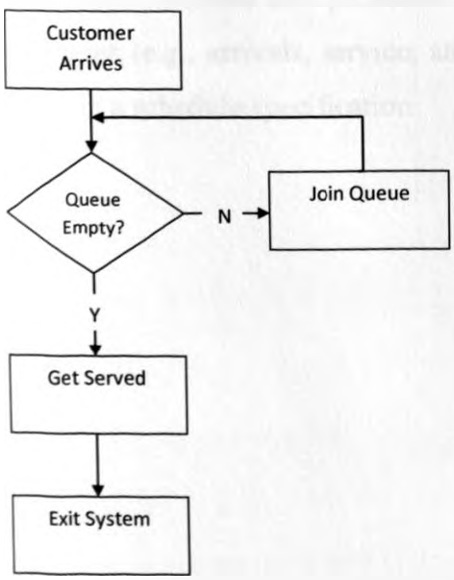


Figure 2.10: Arrival Event

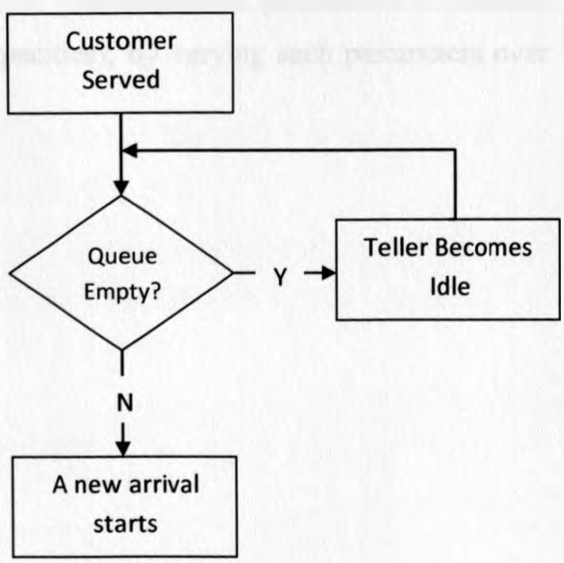


Figure 2.11: Departure Event

- A customer walks into the queue i.e. an **arrival** occurs in the queue (fig 2.10)
- The customer is served and leaves i.e. **departure** occurs from the queue (fig 2.11)

In order to incorporate the above two basic events, in the simulation model, we needed a set of variables known as *clocks*, which kept track of the time instances at which an arrival or departure event occurred. For this specific model the CBS kept track of when a transaction was initiated until it was *committed* or terminated. However, an additional clock was needed to show the time it took the Teller to do other tasks related to a particular transaction before embarking to capture the transaction, this include

- Marshaling of cash (sorting into specific denominations) and
- Verifying customer and account details

Modeling Variability

In this stage, we fitted a probabilistic model to empirical time series data (pairs of time and corresponding observations) collected in Data Collection stage. Variability is about the changes that occur in the components of the system as times goes on. It can be predictable or unpredictable. In this research, the customer inter-arrival time was unpredictable. Typical examples included the "rush hour" phenomenon (temporary heavy traffic) or the "ebb hour" phenomenon (temporary light traffic) [Altiok and Melamed 2007]. Thus, a bank operation experienced "rush hour" periods in the morning (customers stopping by on their way to work) and at lunch time (customers using part of their lunch time for banking), while mid-morning hours were "ebb hour" periods.

Arena has facilities that permitted the modeling of time-dependent parameters of random processes (e.g., arrivals, service, and resource capacities), by varying such parameters over time via a schedule specification

Chapter 3 : METHODOLOGY

3.1 Research Design

This problem could have been resolved in various ways such as: -

1. The management may have depended on the subjective and rule of thumb judgment of an experienced manager
2. Deployed a few personnel first, and then add to the head count as and when the existing personnel would be overwhelmed with work
3. Use benchmarking as a way for calculating staffing requirements, especially when the company lacks competent managerial staff to make informed judgments. For example, a new retail shop may look at the staffing levels of its competitor who sells the same product, has the shop of a similar size, and spends the same amount in advertisements. There is no reason to believe why the business would not require any less or more staff than the competitor.

This research nevertheless, proposed a Modeling and Simulation approach which was scientific and realistic to solving the problem.

3.2 Sources of Data

1. The Bank's Core Banking System (CBS) was the primary source of our input data such as: -
    a. Service times
    b. Total transactions processed
2. Waiting times for service was collected by recording the time it takes customers on joining the queue to the time they got at the teller's station
3. The time it takes for customers to fill in the requisite stationery was collected via observation of significant values and the best fit statistical distributions computed

Data Description

The following data was part of our input to the simulation

1. Service times
2. Customers' inter-arrival times
3. Total transactions
4. Time taken to fill requisite stationery
5. Number of Servers (Tellers)

Service time was defined as the elapsed time from when a teller initiated customer service through the teller transaction platform, to when that customer engagement is ended on the platform. This may exclude some customer engagement time before and after a transaction. Although friendly greetings and small-talk are important for maintaining a positive customer relationship, including them within the service time can be misleading. For instance, tellers may be more likely to have extended conversations with customers during less busy periods in the branch, skewing service time statistics. Our goal in analyzing service times was to recommend staffing levels. So it was better to focus on the actual transaction portion of the customer-teller engagement to get an accurate matching of teller performance to staffing needs.

The Bank classified transactions into various types, these include deposits, cashing checks, purchasing money orders, buying a savings bond, and making a credit card payment.

The three most common transaction types from the data were:-

1. Cash payments (withdrawals)
2. Cash receipt (deposits)
3. Banker's cheque Issuance

This accounted for nearly 97% of all transactions as shown in fig 3.1 below .These three transaction types were the major focus of our analysis.

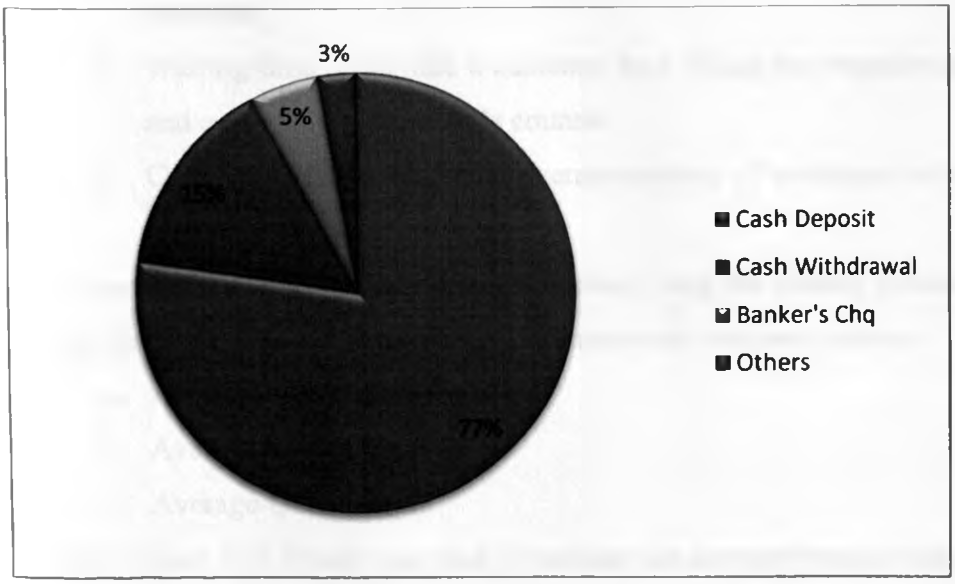


Figure 3.1: Transactions Types Distribution

3.3 Tools

1. **MS Excel** – for data analysis and data clean up so as to have data that can be adapted by ARENA and DTREG and to develop the mathematical/spreadsheet to compute transactions distributions completed on each interval
2. **Modeling and Simulation Software (ARENA form Rockwell)** – provided a less costly approach of experimenting with the model of the real system and also for its great input and output analyzer feature to achieve sufficient level of accuracy. ARENA was also powerful in simulation and modeling of discrete systems.
3. **DTREG** – provided a state of art modeling method and build the decision trees to predict workload/transactions for any given working day of the week (Monday - Saturday)
4. **Visual Basic 6.0** – used to develop an application that computed the teller requirements based on hourly arrivals of the customers

3.4 Procedure

For us to understand the current system, we had to establish its performance using the following steps

1. Collect statistics (performance indicators) this included: -
    a. Total transactions completed at the end of the day computed in one (1) hour intervals
    b. Waiting time – the time a customer took filling the requisite stationery, queuing and eventually at the teller's counter.
    c. Customer Arrival rate – the average number of customers arriving on an hourly interval
2. A simulation of the current system was done using the already collected statistics, from this, the estimated key performance measures were obtained such as :-
    a. Average Queuing time
    b. Average Service time
    c. Average Queue length
3. Spreadsheet (MS Excel) was used to tabulate the average transactions in intervals of one hour each as illustrated in the table (3.1) and graph (3.2) below.

Time	Transactions
8-9am	37
9-10am	55
10-11am	93
11-12noon	50
12-1pm	107
1-2pm	97
2-3pm	115
3-4pm	95
4-5pm	26
Total	**675**

Table 3.1: Transactions competed in one day

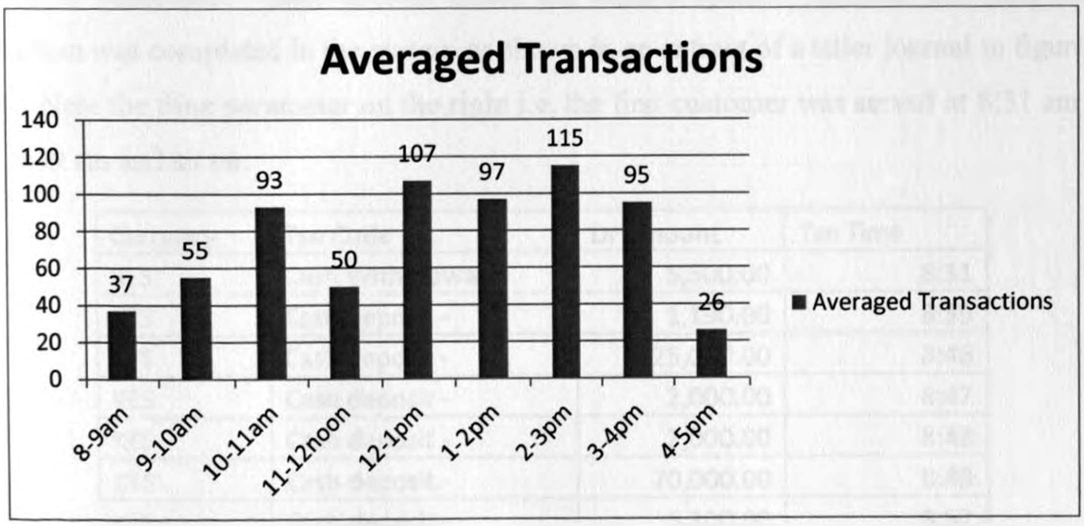


Figure 3.2: Workload distributions for the four Tellers

4. Using the mathematical theory of queues, Little's result, theorem, lemma, law or formula, we computed the estimated queue length on each interval. The theorem states that: The long-term average number of customers in a stable system L is equal to the long-term average effective arrival rate, λ, multiplied by the average time a customer spends in the system, W; or expressed algebraically: $L = \lambda W$ [Little and Graves, 2008].
5. Used the model to calculate Full Time Equivalents (FTEs) needed in each hour interval.
6. DTREG was used to develop decision trees based on the learning data for 2011 and these decision tree models were used to estimate transactions for 2012
7. Developed a user friendly application using VB 6.0 that took in the model and presented to the user the optimal manning requirements

3.4 Input Analysis

The activity of modeling random components is called input analysis. From a methodological viewpoint, it is convenient to temporally decompose input analysis into a sequence of stages, each of which involves a particular modeling activity [Altiok and Melamed, 2007]

1. Data collection
2. Data analysis
3. Time series data modeling
4. Goodness-of-fit testing

Data Collection

To collect **customers' inter-arrival** times the bank's system recorded the time that a transaction was completed in the system as shown in an extract of a teller journal in figure 3.2 below. Note the time parameter on the right i.e. the first customer was served at 8:31 am, the next 8:39 am and so on.

Currency	Txn Code	Dr Amount	Txn Time
KES	Cash Withdrawal	5,500.00	8:31
KES	Cash deposit -	1,190.00	8:39
KES	Cash deposit -	25,000.00	8:46
KES	Cash deposit -	2,000.00	8:47
KES	Cash deposit -	2,000.00	8:48
KES	Cash deposit -	70,000.00	8:49
KES	Cash deposit -	5,300.00	8:52
KES	Cash deposit -	15,000.00	8:53

Table 3.2: Customer Inter-arrivals

From this, we grouped the transactions in intervals of one hour each from start of business (8.30am) to close of business (4.30pm) as shown in the table 3.3 below. This was achieved by amalgamating all that transactions that fell within a given hour and consolidated their total

Time	Total Transactions
8-9am	37
9-10am	55
10-11am	93
11-12noon	50
12-1pm	107
1-2pm	97
2-3pm	115
3-4pm	95
4-5pm	26

Table 3.3: Inter-arrival distribution

To collect **service times** i.e. the time it took to process a transaction on the system (see section 3.2 above). The bank's system captured the time a transaction was completed in the system and thus consecutive timings between preceding transactions gave us the service times as illustrated in an extract of a teller's journal below

Currency	Txn Code	Dr Amount	Txn Time	**Service Times (minutes)**
KES	Cash Withdrawal	5,500.00	8:31	0
KES	Cash deposit -	1,190.00	8:39	8
KES	Cash deposit -	25,000.00	8:46	7
KES	Cash deposit -	2,000.00	8:47	1
KES	Cash deposit -	2,000.00	8:48	1
KES	Cash deposit -	70,000.00	8:49	1
KES	Cash deposit -	5,300.00	8:52	3
KES	Cash deposit -	15,000.00	8:53	1
KES	Vault to Teller	300,000.00	8:57	4
KES	Cash deposit -	11,000.00	8:58	1
KES	Cash deposit -	50,000.00	9:05	7
KES	Cheque Deposit	20,000.00	9:17	12
KES	Cash deposit -	24,119.00	9:21	4

Table 3.4: **Teller Journal**

We executed queries on the bank's system to get the **total transactions** of a day, and had the output in a spreadsheet form for ease of manipulation and analysis. A sample output is shown in Table 3.5 below

XYZ BANK LTD			
BRANCH PERFORMACE REPORT BY USER			
Run Date	04-Jan-11		
Total Branch Txns	675		
Branch	Z Street		
User Id	User Name	FT Txns	TT Txn(s)
KE10923	DOROTHY		146
KE11064	HILDA	3	148
KE11096	NANCY	3	206
KE8901	KIRIMI	3	175

Table 3.5: **Branch Total Activity**

To get the **time taken to fill requisite stationery** we will observe and record time taken by the customers

Data Analysis

The analysis stage often involved the computation of various empirical statistics from the collected data, including

- Statistics related to moments (mean, standard deviation, coefficient of variation, etc.)
- Statistics related to distributions (histograms)
- Statistics related to temporal dependence (autocorrelations within an empirical time series, or cross-correlations among two or more distinct time series) [Altiok and Melamed, 2007].

Descriptive model in the form of tables and histograms was used to organize and summarize the input data being analyzed.

Theoretical model, which sought to test whether or not the phenomenon (arrival pattern) being observed, conformed to various mathematical or statistical theories e.g. a beta or exponential distribution. This was addressed by ARENA's input analyzer.

Stationery Time Analysis

We eventually fitted a statistical distribution using ARENA's input analyzer to get the service time input variable to our model as shown below with the best fitted outcome being a Weibull distribution with the following parameters

Distribution Summary

Distribution: Weibull
Expression: 0.5 + WEIB(3.77, 0.961)
Square Error: 0.002412

Chi Square Test
Number of intervals = 12
Degrees of freedom = 9
Test Statistic = 26.6
Corresponding p-value < 0.005

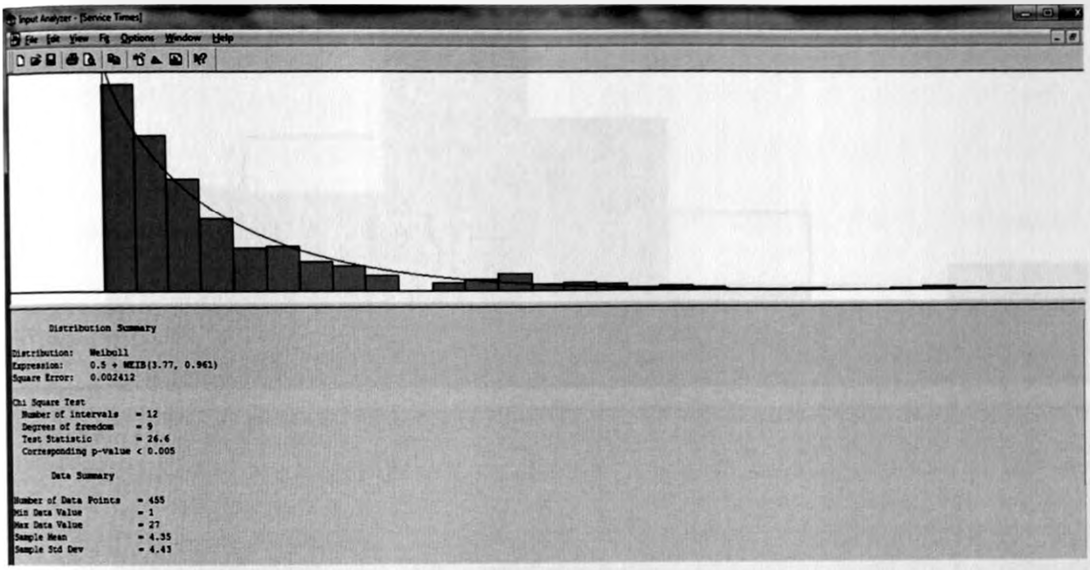


Figure 3.3: Time taken to fill Stationery

Service Time Analysis

Distribution Summary

Distribution:	Poisson
Expression:	POIS(3.46)
Square Error:	0.011508

Chi Square Test

Number of intervals	= 7
Degrees of freedom	= 5
Test Statistic	= 87.4
Corresponding p-value	< 0.005

Data Summary

Number of Data Points	= 675
Min Data Value	= 1
Max Data Value	= 7
Sample Mean	= 3.46
Sample Std Dev	= 1.75

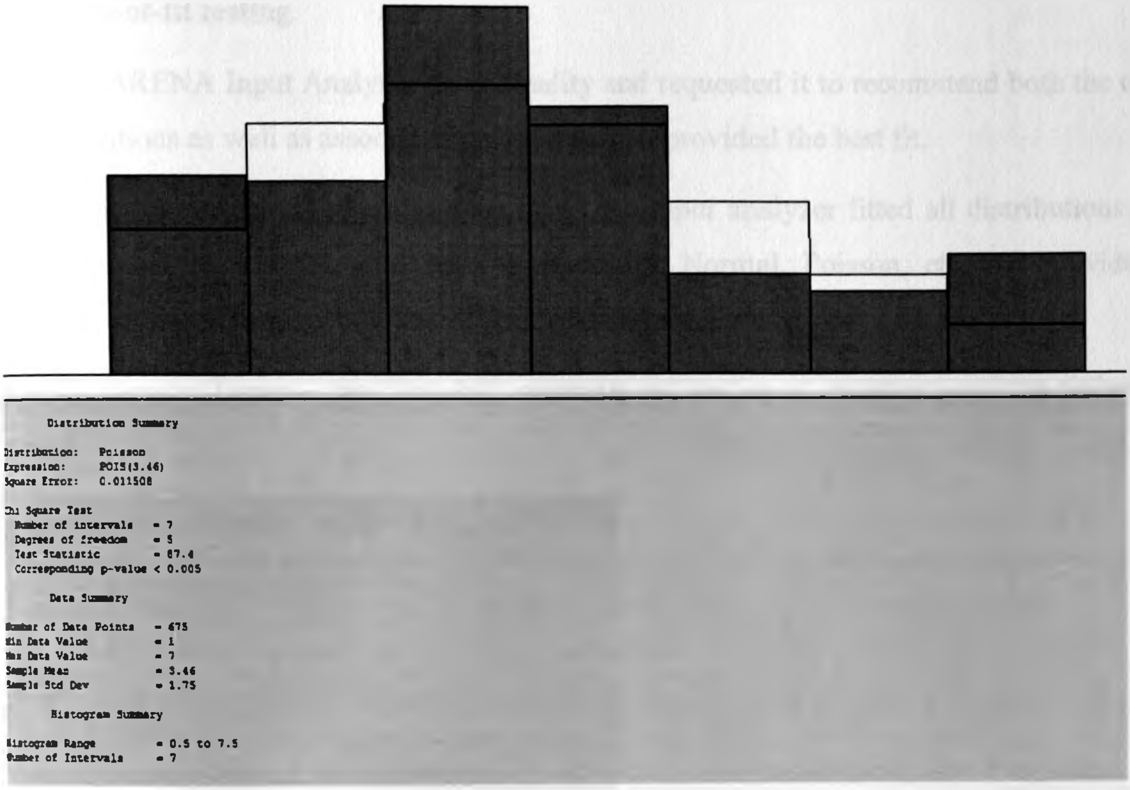


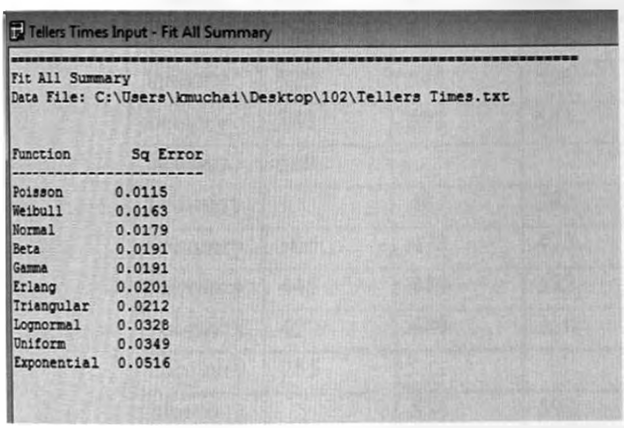
Figure 3.4: Time taken to serve customer

Goodness-of-fit testing

We used ARENA Input Analyzer functionality and requested it to recommend both the class of distributions as well as associated parameters that provided the best fit.

For us to pick on one distribution over other, the input analyzer fitted all distributions that ARENA supports e.g. Gamma, Beta, Exponential, Normal, Poisson, etc. and provided a summary of the SQ errors for all as illustrated in figure 3.5

For the Service Times – the "Fit All" summary of the input analyzer gave the following output



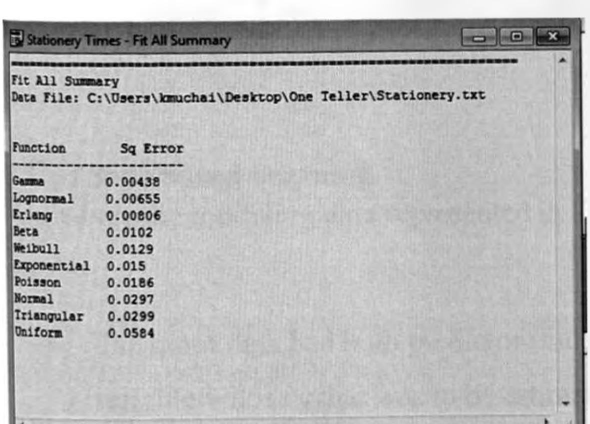
Tellers Times Input - Fit All Summary

Fit All Summary
Data File: C:\Users\kmuchai\Desktop\102\Tellers Times.txt

Function	Sq Error
Poisson	0.0115
Weibull	0.0163
Normal	0.0179
Beta	0.0191
Gamma	0.0191
Erlang	0.0201
Triangular	0.0212
Lognormal	0.0328
Uniform	0.0349
Exponential	0.0516

Figure 3.5: Goodness-of-fit Teller's point

Fort the time taken to fill stationery – the "Fit All" Summary gave the following output



Stationery Times - Fit All Summary

Fit All Summary
Data File: C:\Users\kmuchai\Desktop\One Teller\Stationery.txt

Function	Sq Error
Gamma	0.00438
Lognormal	0.00655
Erlang	0.00806
Beta	0.0102
Weibull	0.0129
Exponential	0.015
Poisson	0.0186
Normal	0.0297
Triangular	0.0299
Uniform	0.0584

Figure 3.6: Goodness-of-fit Stationery point

The Square Error field appears in each summary of the distribution fit. It provides an important measure, of the goodness-of-fit of a distribution to an empirical data set

3.5 Workload Prediction

3.5.1 Data Mining Process

Being the process of extracting useful information from a set of data values [Philip H. Sherrod, 2003 -2012]

Using transactional data for 2011 where a total of 148,429 transactions were completed for the branch for 304 working days. This data was represented monthly (January – December) for each day of the week (Monday – Saturday) as shown in the table extract below.

Month	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
January	533	675	689	702	627	356
January	649	681	634	630	663	372
January	629	544	556	567	515	327
January	585	577	543	472	515	299
January	659					
February		487	582	576	481	297
February	505	478	473	530	460	280
February	445	414	522	427	409	240
February	492	489	350	421	529	309
February	285					
March		576	592	643	589	368
March	508	481	520	479	440	290
March	514	504	445	463	441	225
March	419	395	403	414	412	252
March	484	473	418	449		

Table 3.6: Daily Workload

3.5.2 Supervised Learning

To analyze the modeling data represented in Table 3.6 above, we used supervised learning as:-

- The input data had both predictor (independent) variables and target (dependent) variable whose value was to be estimated and
- The goal was to predict the value of some variable

For example, to predict the value for Saturday (target), all other variables (month, Monday – Friday) would form the predictor variables

3.5.3 Variable Types

Our data set had: -

1. Continuous variables (i.e. with numerical values such as 102,555, etc.) in our dataset the variables are days of the week
2. Categorical variables (i.e. with values that function as labels such as January, February, etc.) in our dataset we have the variable as month

3.5.4 Decision Tree

This is a logical model represented as a binary (two-way split) tree that shows how the value of a target variable can be predicted by using the values of a set of predictor variables [Philip H. Sherrod (2003 -2012)]

Since our goal was to learn the data and predict some value, we developed decision trees using historical transaction data. DTREG used this data to learn how the value of the target variable was related to the value of predictor variables. To improve on accuracy level, we modeled the whole set for 2011. To test the accuracy of the estimated values, we used actual data of January 2012 and compared with the estimate outcome as shown in the results section. DTREG performed complex analysis on this data and built decision trees that modeled the data

Why decision trees?

1. They are easy to interpret even for non-technical people
2. It's possible to predict target values for specific cases where only the predictor variables are known

Regression v/s Classification Model

The target variable dictated whether continuous or categorical model was generated. If the target variable was continuous, DTREG would generate a regression model and if categorical, a classification model was generated. Since out target (number of transactions) was continuous, we generated a regression model.

Setting the variables to either target or predictor is illustrated in figure 3.7 below

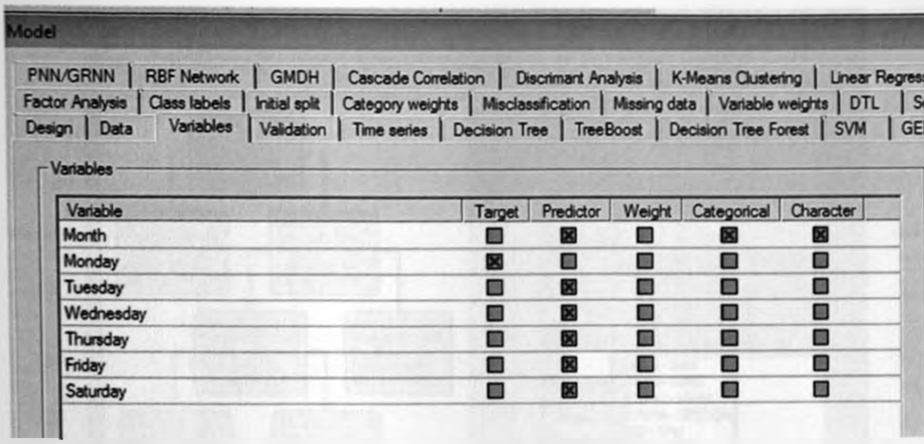


Figure 3.7: Predictor & Target variables

3.5.5 Application Design

The Teller Resource system was developed with Microsoft Visual Basic 6.0 which provided an easy to use GUI that the user could click on the predictions of each individual day and the respective decision tree is displayed.

Figure 4.3 shows the first window that welcomes the user to the system where options of viewing the decision tree models are activated by the command buttons as highlighted

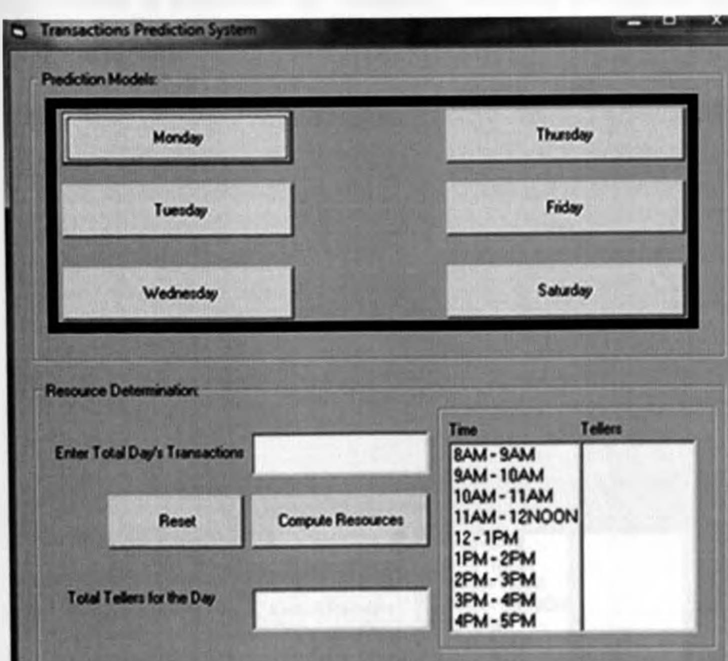


Figure 3.8: Interface Design

Figure 4.4 shows the model for "Friday" that is activated by the command button labeled "Friday" as highlighted below.

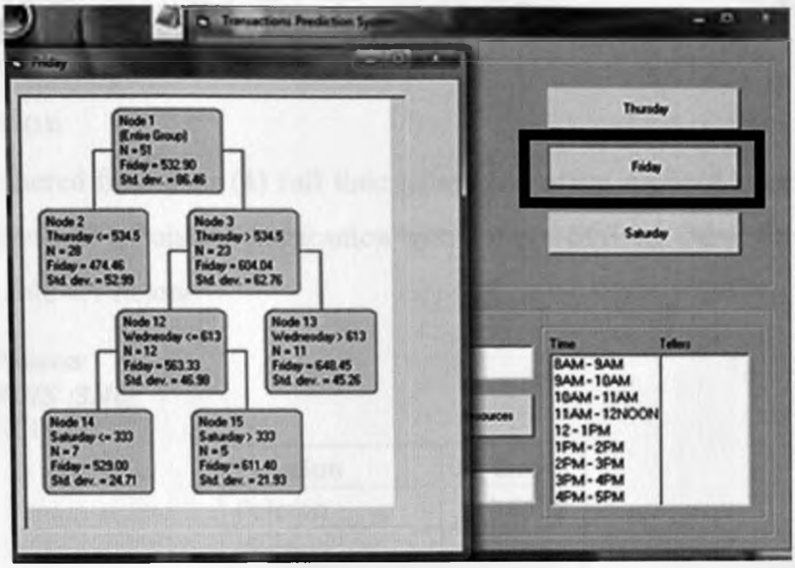


Figure 3.9: Friday's Prediction Tree

Figure 4.5 demonstrates how the system computes the manning levels for each hour 8 – 9am, 9 – 10 am and so on based on the value of the total transaction that is entered on the text box provided. It eventually sums up the total tellers that should be dedicated for that day.

The total transaction value is achieved by 'walking' the tree model of the day sought based on the input/predictor variables.

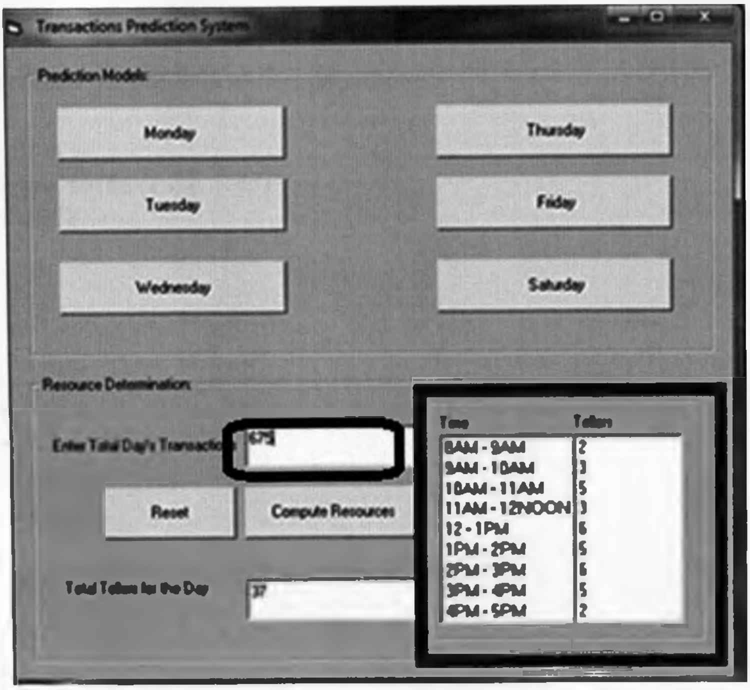


Figure 3.10 Tellers Schedule

Chapter 4 : RESULTS

4.1 Simulation

The data was gathered from four (4) full time tellers and using ARENA's input analyzer, the best fitted function was a Poisson distribution with a mean of 3.46. Other fitted functions are summarized in table 4.1 below

Distribution: Poisson
Expression: ***POIS (3.46)***
Square Error: 0.011508

Function	Sq Error
Poisson	0.0115
Weibull	0.0163
Normal	0.0179
Beta	0.0191
Gamma	0.0191
Erlang	0.0201
Triangular	0.0212
Lognormal	0.0328
Uniform	0.0349
Exponential	0.0516

Table 4.1: Service Distribution

In the same breath, the time taken by customers to fill up pre-requisite stationery was fitted to a Gamma distribution with the following parameters

Distribution: Gamma
Expression: ***0.5 + GAMM (1.84, 1.63)***
Square Error: 0.004378

4.2 The Proposed Manning Model - Objective 1

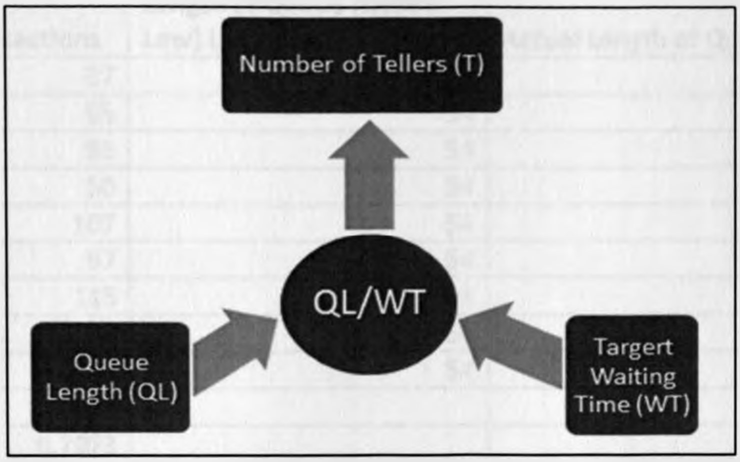


Figure 4.1: Proposed Manning Model

Where:-

Queue Length (QL) – was defined by Little's law

Target waiting time (WT) – the minimum duration that was set by Bank's management that a customer should wait for service in minutes, this was set to 15mins

Number of Tellers (T) – the calculated tellers/resources by the model

i. From the collected statistics, the current average waiting time was 0.7073 hrs. i.e. 42mins
ii. The average arrival rate was75 customers per hour
iii. We used Little's Law and assumed that this was a stable system, i.e. the rate at which people enter the bank is the rate at which they exit as well, the outcome of this is shown in table 4.2
iv. On the assumption that the tellers processing speeds was the same, the processing time was equal and the system was stable. The length of the queue (L) was given by (λ =75) multiplied by (W= 07073).
v. Comparing these results with the actual arrivals will generate the actual length of the queue on each hourly interval as shown in table 4.2, column four emphasized below.

Queue Length Estimate				
Time	Transactions	Length of Queue (Little's Law) L= λ W	Actual Length of Queue	Queue Status
8-9am	37	54	-17	No Queue
9-10am	55	54	1	Queue
10-11am	93	54	39	Queue
11-12noon	50	54	-4	No Queue
12-1pm	107	54	53	Queue
1-2pm	97	54	43	Queue
2-3pm	115	54	61	Queue
3-4pm	95	54	41	Queue
4-5pm	26	54	-28	No Queue
Total Txns	**675**			
W =	**0.7073**			
λ =	**75**			

Table 4.2: Queue Length Estimate

vi. During the 8am – 9am interval where we had 37 arrivals, there was no queue hence the negative value (-17). The queue grows for all other positive values say the interval between 10 – 11am with an estimate of 39 customers on the queue.

vii. We set the targeted waiting time (WT) to 15 min

viii. Applying the model shown in figure 4.1 generates the resources as presented in table 4.3

Modeled Resources			
Time	Transactions (λ)	Length of Queue (Little's Law) L= λ W	Model Resources (Tellers)=Queue Length/WT
8-9am	37	27	2
9-10am	55	39	3
10-11am	93	66	5
11-12noon	50	36	3
12-1pm	107	76	6
1-2pm	97	69	5
2-3pm	115	82	6
3-4pm	95	67	5
4-5pm	26	19	2
Total Txns	**675**		
W =	**0.7073**		

Table 4.3: Resources Calculation

4.3 Transaction Estimation – Objective 2

Using the Decision Tree to Predict a Target Variable (Scoring)

Let's predict (score) the target variable (Thursday) in the sample data below (table 4.4) using the regression tree model generated for Thursday (figure 4.2) and all other inputs as predictor variables

Predictor	Predictor	Predictor	Predictor	**Target**	Predictor	Predictor
Month	Monday	Tuesday	Wednesday	**Thursday**	Friday	Saturday
March	508	471	444	**487**	428	289

Table 4.4: Scoring Example

We begin with the root node (node 1). Then decide whether to go to the left or right node based on the value of the splitting variable (Wednesday). In our case, the value for Wednesday = 444 which leads us to node2. Since node 2 is not a terminal node we continue to use the splitting variable, i.e. "Saturday"<=262.5. Our predictor value is 289 and that lead us to node 5, which is still not a terminal node. The last splitting variable is Month, in our case, Month = "March" and this leads us to node 8 which is a terminal node whose value for Thursday = 481

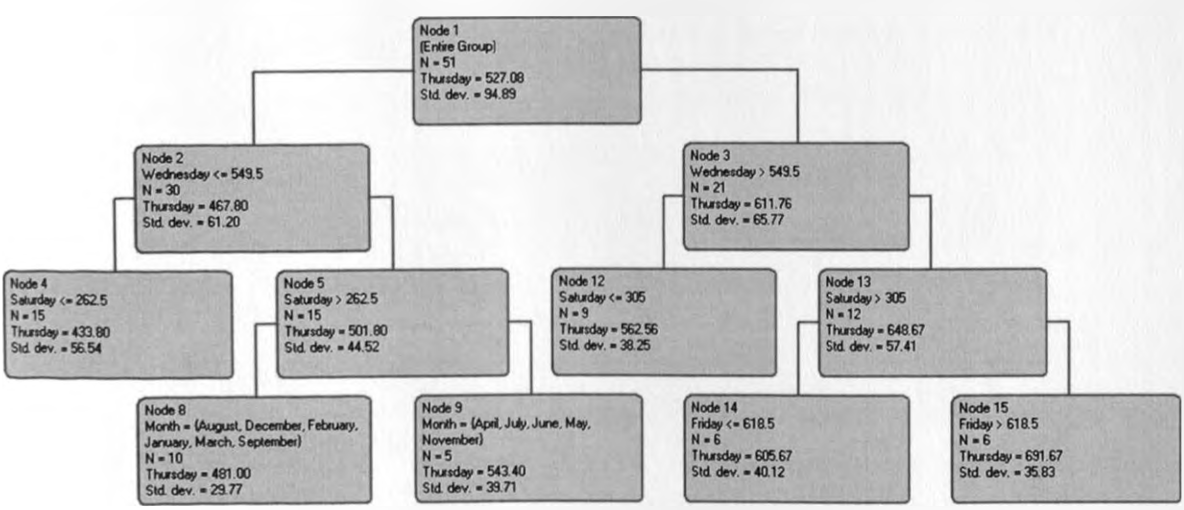


Figure 4.2: Thursday Regression Tree Model

Accuracy Testing

To test the accuracy of the prediction, we used predicted values of the model (table 4.5) and tested against actual values for January 2012 (table 4.6).

Estimated Values Results						
Month	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**
Jan	746	547	616	691	611	375
Jan	534	547	541	481	474	283
Jan	613	454	576	481	611	315
Total	**1893**	**1548**	**1733**	**1653**	**1696**	**973**

Table 4.5: Estimated values for Jan 2012

Actual Data for Jan 2012						
Month	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**
Jan	803	757	592	620	714	410
Jan	634	626	434	436	542	333
Jan	525	584	530	595	562	431
Total	**1,962**	**1,967**	**1,556**	**1,651**	**1,818**	**1,174**

Table 4.6: Actual values for Jan 2012

The estimation accuracy achieved was 90% as shown in table 4.7

Accuracy Measure						
Residual Value	**69**	**419**	**-177**	**-2**	**122**	**201**
Daily Accuracy	96%	79%	90%	100%	93%	83%
Average Accuracy	**90%**					

Table 4.7: Accuracy testing

4.4 Waiting Time reduction – Objective 3

The model reduced the total time spent by customers with an approximate 60% from 36 to 20 min. The proposed model scheduled the tellers with intervals of one hour each and was driven by the demand/customer arrivals. Simulation results for the current system with four (4) full time tellers, gave the results in table 4.8, while the proposed model results are shown in table 4.9

Current System

Current System		
Performance Measure	**Average Time (hrs)**	**Average Time (mins)**
Service & Stationery Time	0.114	7
Queuing Time	0.5933	36
Total Time	**0.7073**	**43**

Table 4.8: Waiting time Current System

Proposed System

Proposed System		
Performance Measure	**Average Time (hrs)**	**Average Time (mins)**
Service & Stationery Time	0.115	7
Queuing Time	0.3344	20
Total Time	**0.4494**	**27**

Table 4.9: Waiting time Proposed System

Queue Performance

The performance of the queue was based on: -

- Duration spent on the queue
- Number of people waiting

The duration spent by customers is shown in tables 4.8 and 4.9 for both the current and proposed systems respectively and the average number of people waiting is shown in table 4.10.

System	Avg. Number Waiting
Current	36
Proposed	21

Table 4.10: People waiting

4.5 Results Analysis

The model's performance is based on the assumption that there is no physical limitation in terms of the teller cubicles as transactions can demand a higher number say from lows of two (2) to highs of eight (8) on some hours during some busy days. This model minimises the waiting time with a significant margin which was the core motivator to this study. The model opens up an the human resources domain advocating for a more scientific approach in deploying the same, the advantage is that one can get to measure its performance and compare with the traditional static set-up

While assessing queue performance, many factors came into play such as the configuration, the general assumption was that the tellers were universal and could perform any transaction and thus the bank patrons were served at any teller's point.

Section 4.2 details how the model was arrived at with the backing of Little's law which has been scientifically proven. Little's law performance is based on the drivers i.e. the waiting time is influenced by the service rate and thus the more the number of servers the higher the service rate, however and optimal number has to be arrived at so as to have some equilibrium in terms of cost v/s revenue.

Section 4.3 details how estimation was being done via decision trees. This provided an easy to use approach method for scoring even for novice users. To improve on the accuracy, more data should be used. This research used data for 2011 thus continuous improvement should be done by incorporating new data and by so doing improve on the accuracy. Knowing/estimating the number of patrons to expect beforehand provides for planning for the future and deploying the appropriate number thus the question of long waits is addressed.

Section 4.4 expounds on how the waiting times was reduced. This is a business booster that most service industries would want to incorporate. Time is a valuable asset and a minute saved at the banking hall could mean having ample time to attend to other errands as people will always be on the move.

Chapter 5 : DISCUSSIONS

The study of Queues was instigated by Erlang, a Danish Engineer whose focus area was call centers. His works has since extended to other areas in life where queues are the order of the day. Say for example in the service industry like Banks, Supermarkets, etc. This brings us to the question, why do queues form? This question has two responses i.e.

1. A queue forms when the demand for a service facility exceeds the capacity of that facility i.e. the customer do not get served immediately upon arrival but must wait. And
2. A queue forms when some customers wait when the number of customers requiring service exceeds the number of service facilities, some service facilities stand idle when the total number of service facilities exceeds the number of customers requiring service.

Addressing the responses above motivated this study. First is the demand for service (customers) exceeding the capacity of the facility (tellers). This ultimately creates a bottleneck and the patrons have to wait. So how would a lay person address this, the top of head solution would be to increase the number of teller say there are three tellers to serve the customers a quick solution would be to add say three more and thus have six tellers to server the customers and thus there will be no customer queue. However this will yield another queue as highlighted in 2 above where the servers stay idle i.e. the service facility will exceed the number of customers requiring this service. Another restriction to this quick fix is the physical constraint in that the service facility could be made such that it can only accommodate a fixed number of tellers. Tabling this solution to cost conscious proponents will face fierce opposition as deploying idle resources impacts negatively to the organization's revenue.

So as to understand a queue system, one need to assess what affects its performance is which is dictated by:-

1. Input Source
2. Service System
3. Queuing Discipline

The Input Source is categorized as Static or Dynamic. Most queuing systems are founded on Static arrival process where the probability of arrivals is described as the number of customers arriving per unit of time. This study however proposed a Dynamic approach where the Service Rate was determined by both the service facility (Tellers) as well as the arrivals (Customers). The service facility adjusted its capacity to match the changes in demand intensity by varying the staffing levels at different timings of the day.

The model, i.e. $\boldsymbol{Number\ of\ Tellers(T)} = \frac{\boldsymbol{Queue\ Length(QL)}}{\boldsymbol{Target\ Waiting\ Time\ (WT)}}$

Where QL was defined by Little's law as L= λ W.

- L – Length of the queue
- λ – Average arrival rate. This was obtained from the hourly arrivals as the customer transactions were clustered into arrivals of one hour each
- W – Average waiting time. This was obtained from the simulation runs on the collected data

From λ and W above, L was obtained which translated to QL in the model. Dividing this with the value of the Target Waiting Time (WT) whose value was set to 15min; this resulted into a value that represented the number of tellers that would adequately manage the customer arrivals for that given hour. This was replicated throughout the day to result in the dynamic teller schedule for that day.

To estimate the work volumes/transactions that the branch would expect, supervised learning was used via DTREG which performed the analysis of the input data and built decision trees that conformed to the data. Decision trees were preferred since our goal was to learn the data and predict some value. To get a predicted value one needed to 'walk' the tree based on the predictor variables until you arrive at a terminal node with translated to the target variable. The accuracy of this mode of estimation was arrived at 90% as detailed in section 4.3

This model effectively reduced the waiting time by 60% this was obtained from running simulations of the current system with four full time tellers and the modeled dynamic resources as expounded in section 4.4

An application was developed using VB6.0 to give a user friendly interface to the intended user of such model, for them to just key-in the predicted value by the decision tree and it automatically generated the resources schedule for the day. This enabled the Manager in-charge of resource deployment to plan in advance the workforce needed to meet the day's customers, which translates to value-added customer service.

CONCLUSION AND FUTURE WORK

The objectives of this study were to:-

1. Develop a model that can be used to determine teller requirements
2. Predict transaction volumes for each day of a week.
3. Reduce the waiting times for service

Our efforts produced a model developed from the queuing theory and Little law principles, illustrating the inadequacy of fixed full time manning system of the teller service stations and highlighted an optimal and effective solution of the scheduling system that would meet the bank's desired service levels.

The literature studied was from scientific literature with key findings such as it is the ideal combination of using simulation models next to queuing theory/models to obtain best results. The simulation model can be used to perform what-if scenarios, because of high flexibility. Apart from the fact that simulation provides the possibility to model more complex processes. The queuing model was used to approximate the system performance measures. Green, Kolesar and Whitt in their study demonstrated the link between the queue and the capacity or resources which is evident in this study as the Resources/tellers were computed on the basis of the arrival rate of the customers

When following an event scheduling perspective, a model developer must define the model logic and system state changes that occur whenever any event occurs. For example, a customer process at a bank consists of an arrival event (to the lobby, perhaps), joining and waiting in a queue (a delay), a service time by a teller, and finally a service completion event. In terms of concepts discussed earlier, the service time is an activity and the teller is a resource.

The study was based on several assumptions that were made on the onset, a major assumption was that of assuming that all tellers are Universal i.e. are equal in skills and processing speeds. Banks have come up with generalists and specialists arrangements to hasten the service to customers and reduce on waiting times. Generalists are agents that can handle all types of customers, and specialists are agent that can handle only one type of customer [Stolletz (2003)]. With this in mind future work could be developed around this area to translate the model to fit the real world and assess the effect on the waiting time for service.

The estimation aspect proposed by this model, seeks to address the shortcoming of long-term planning that is evident in majority of scheduling systems with an accuracy of 90%

REFERENCES

Abratt, R., & Russell, J. (1992). *The International Journal of Bank Marketing.* South Africa.

David Cohen, Christopher Gan, Hua Hwa Au Yong, & Esther Choong. (2006). *Customer Satisfaction: A Study of Bank Customer Retention.* New Zealand.

Donald Hammond, & Sathi Mahesh. (1995). A Simulation and Analysis of Bank Teller Manning. *Proceedings of the 1995 Winter Simulation Conference.*

Efraim Turban, & Jay E. Aronson. (1997). *Decision Support Systems and Intelligent Systems.*

Henderson W.B. , & Berry W. L. . (1977). *Determining Optimal Shift Schedules for Telephone Traffic Exchange Operators.*

Hull, L. (2002). *Foreign-owned Banks: Implications for New Zealand's Financial Stability.* New Zealand.

Ivo Adan, & Jacques Resing. (2002). *Queuing Theory.*

John D.C, Little, & Stephen C. Graves. (2008). *Little's Law.*

Kaynak, E., & Kucukemiroglu O. (1999). Bank and Product Selection. *The International Journal of Bank Marketing*, Vol. 17, No. 11. – pp. 5-19.

Linda V. Green, Peter J. Kolesar, & Ward Whitt. (2007). *Coping with Time-Varying Demand When Setting Staffing Requirements for a Service System.*

Mark .J. Stegeman, & Monique. H. Jansen-Vullers. (2005). *Determining Optimal Staffing Levels in Multiple Skill Inbound Call Centers.*

Philip H. Sherrod. (2003 - 2012). *DTREG Predictive Modeling Software.*

Pihl, W. H., & Wambay M. L. . (1990). The Leading regionals focus on productivity. *Bank Management*, 66: 52-55.

Robinson Stewart. (2004). *Simulation: The Practice of Model Development and Use.*

Scott Snell, & George Bohlander. (2007). *Human Resource Management (International Student Edition).*

Tayfur Altiok, & Benjamin Melamed. (2007). *Simulation Modeling and Analysis with Arena.*

The Oxford Dictionary . (n.d.).

Travis Cogdill, & Michael Monticino. (2007). *Analysis of Teller Service Times in Retail Banks.*

APPENDICES

1. **Project Schedule**

MSc. Computer Science Project Schedule

ID	Milestone Title	Start	Finish	Duration	Mar 2012	Apr 2012	May 2012	Jun 2012	Jul 2012
					4/3	1/4 8/4	6/5	3/6	1/7 8/7
1	Proposal Presentations	05/03/2012	16/03/2012	2w					
2	Progress Presentations	11/06/2012	22/06/2012	2w					
3	Final Presentations	23/07/2012	03/08/2012	2w					

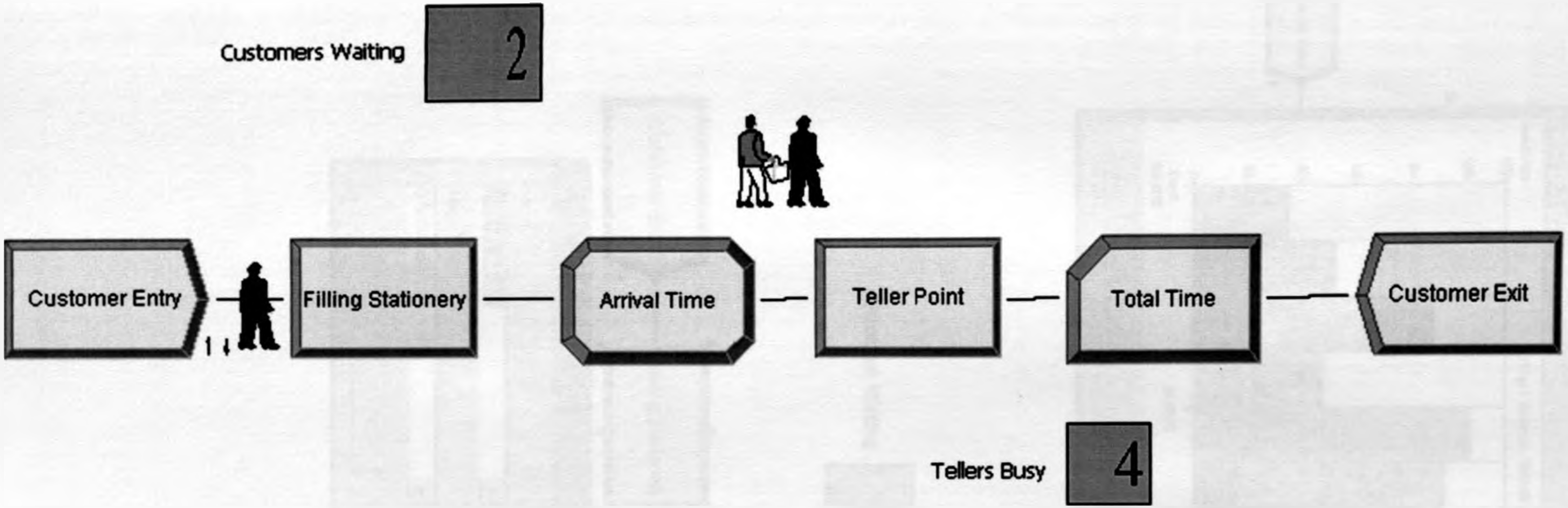


Figure 0.1: The Simulation in progress

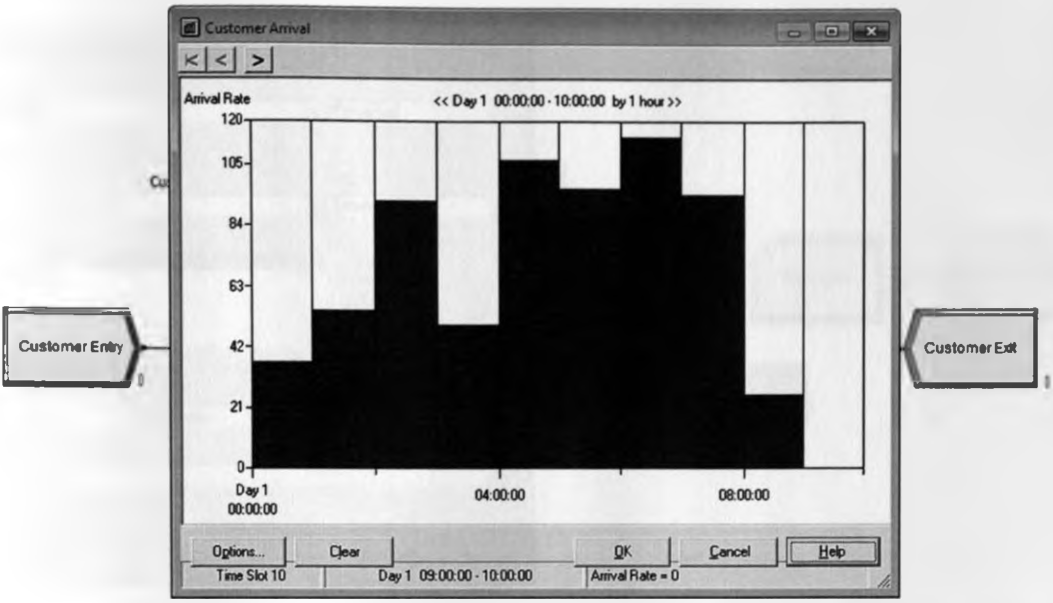


Figure 0.2: The Arrivals Distribution

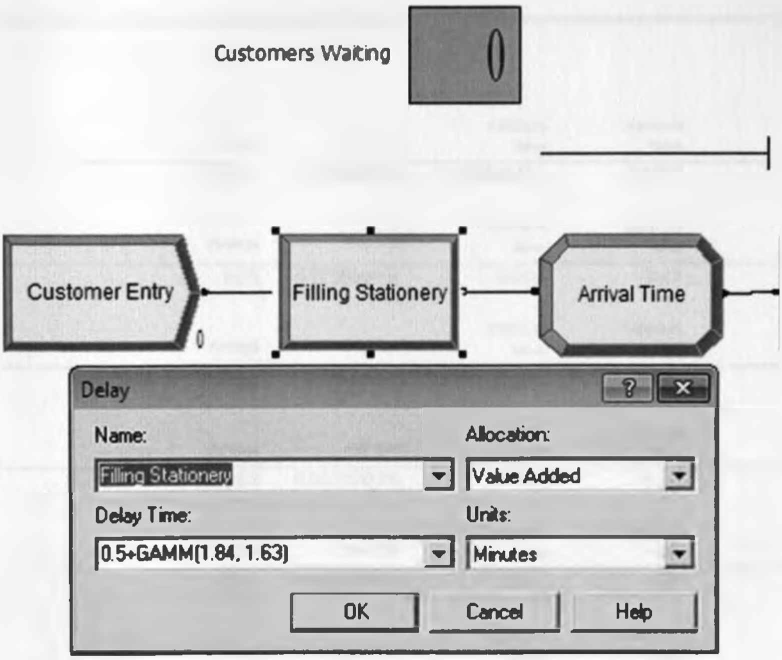


Figure 0.3: The Stationery Module

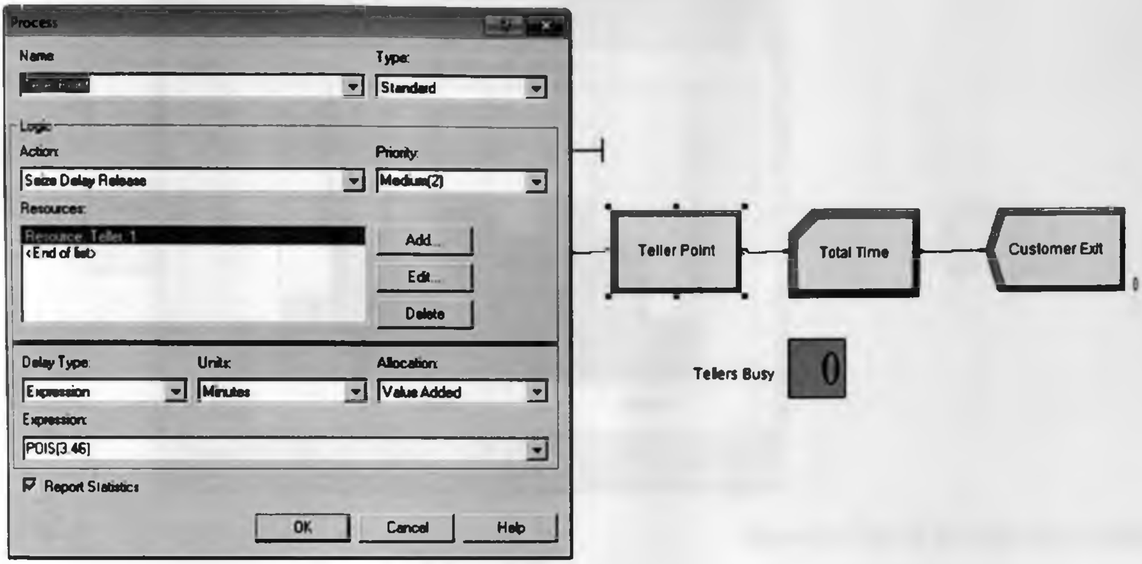


Figure 0.4: The Teller Service Module

Entity

Time

VA Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.1140	0.003842282	0.01584412	0.3091

NVA Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.00	0.000000000	0.00	0.00

Wait Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.5933	(Correlated)	0.00	1.8621

Transfer Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.00	0.000000000	0.00	0.00

Other Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.00	0.000000000	0.00	0.00

Total Time	Average	Half Width	Minimum Value	Maximum Value
Customer	0.7073	(Correlated)	0.02708480	2.1082

Figure 0.5: Sample Report

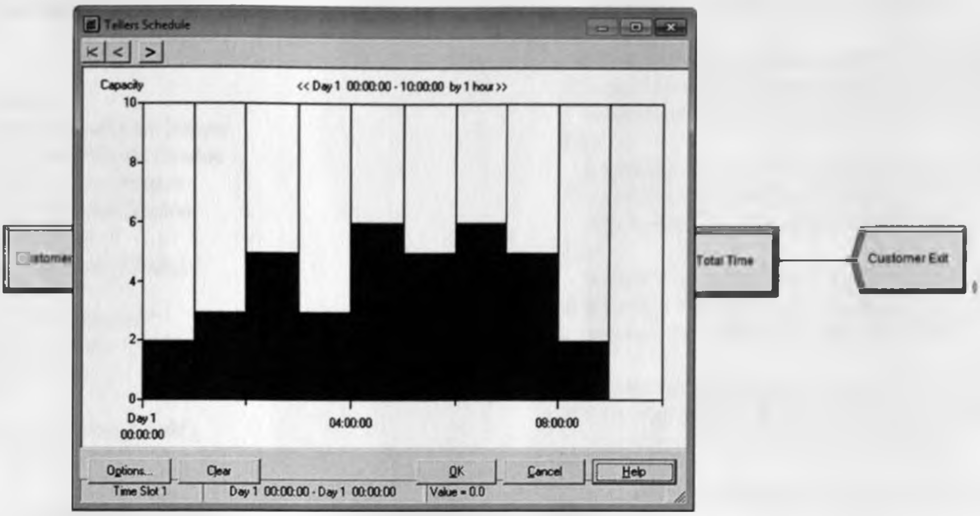


Figure 0.6: The Teller Schedule Module

3. DTREG Objects

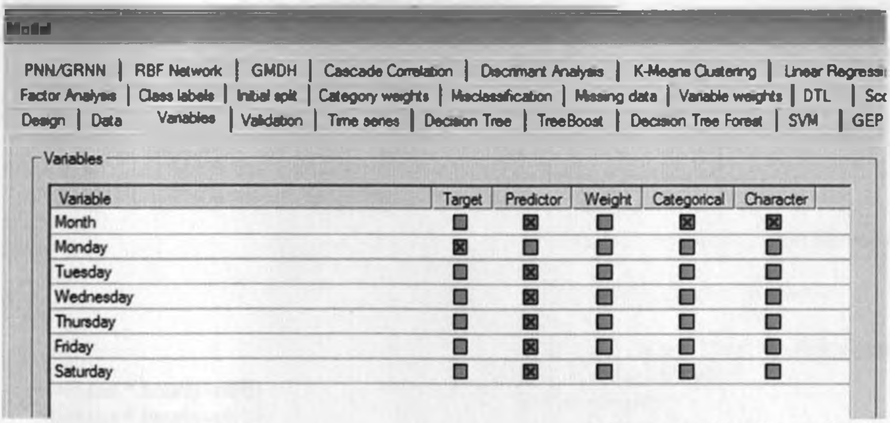


Figure 0.7: Predictor & Target variables

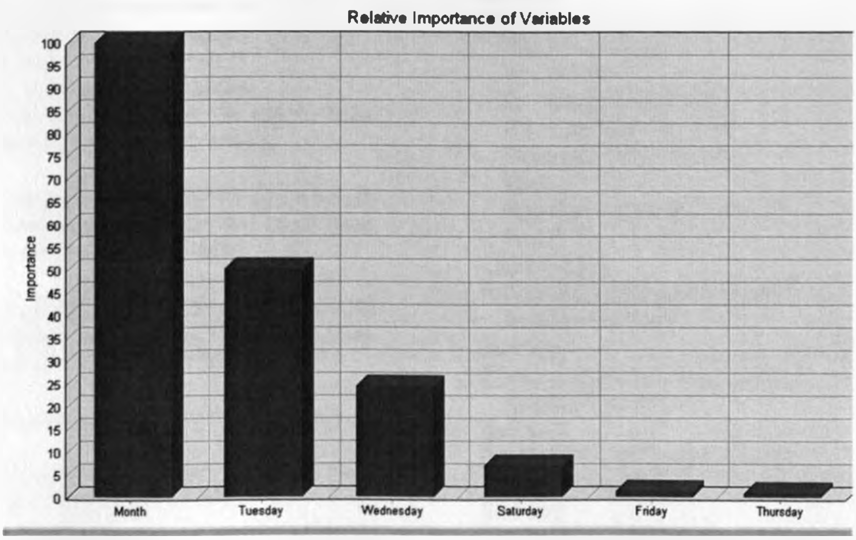


Figure 0.8: Predictor Variable importance

4. Sample Code

```
Option Explicit
  Private hourlytxns(8) As Integer
  Private hourlyvol(8) As Double
  Private totaltxns As Integer
  Private tellers(8) As Double
  Dim L As Double
Private Sub cmdClear_Click()
  txtTxns.Text = ""
  txtTotalTellers.Text = ""
  HourlyTellers.Clear
  txtTxns.SetFocus
End Sub
Private Sub cmdFriday_Click()
  frmFriday.Show
End Sub
Private Sub cmdMonday_Click()
  frmMonday.Show
End Sub
Private Sub cmdResources_Click()
  HourlyTellers.Clear
  txtTxns.SetFocus
  Dim hours As Integer
  Dim W As Double
  totaltxns = Val(txtTxns.Text)
  W = 0.7073
  If IsNumeric(totaltxns) And totaltxns > 250 And
totaltxns < 1000 Then
  hourlyvol(0) = 0.0548
  hourlyvol(1) = 0.0815
  hourlyvol(2) = 0.1378
  hourlyvol(3) = 0.0741
  hourlyvol(4) = 0.1585
  hourlyvol(5) = 0.1437
  hourlyvol(6) = 0.1704
  hourlyvol(7) = 0.1407
  hourlyvol(8) = 0.0385
  hourlytxns(0) = totaltxns * hourlyvol(0)
  hourlytxns(1) = totaltxns * hourlyvol(1)
  hourlytxns(2) = totaltxns * hourlyvol(2)
  hourlytxns(3) = totaltxns * hourlyvol(3)
  hourlytxns(4) = totaltxns * hourlyvol(4)
  hourlytxns(5) = totaltxns * hourlyvol(5)
  hourlytxns(6) = totaltxns * hourlyvol(6)
  hourlytxns(7) = totaltxns * hourlyvol(7)
  hourlytxns(8) = totaltxns * hourlyvol(8)
  If tellers(0) = Round(hourlytxns(0) * W / 15, 0) Then
    tellers(0) = Round(hourlytxns(0) * W)
  Else
    tellers(0) = Round(hourlytxns(0) * W / 15 + 0.5, 0)
  If tellers(1) = Round(hourlytxns(1) * W / 15, 0) Then
  tellers(1) = Round(hourlytxns(1) * W)
  Else
    tellers(1) = Round(hourlytxns(1) * W / 15 + 0.5, 0)
  If tellers(2) = Round(hourlytxns(2) * W / 15, 0) Then
  tellers(2) = Round(hourlytxns(2) * W)
  Else
    tellers(2) = Round(hourlytxns(2) * W / 15 + 0.5, 0)

  If tellers(3) = Round(hourlytxns(3) * W / 15, 0) Then
  tellers(3) = Round(hourlytxns(3) * W)
  Else
    tellers(3) = Round(hourlytxns(3) * W / 15 + 0.5, 0)
  If tellers(4) = Round(hourlytxns(4) * W / 15, 0) Then
    tellers(4) = Round(hourlytxns(4) * W)
  Else
    tellers(4) = Round(hourlytxns(4) * W / 15 + 0.5, 0)
  If tellers(5) = Round(hourlytxns(5) * W / 15, 0) Then
    tellers(5) = Round(hourlytxns(5) * W)
  Else
    tellers(5) = Round(hourlytxns(5) * W / 15 + 0.5, 0)
  If tellers(6) = Round(hourlytxns(6) * W / 15, 0) Then
    tellers(6) = Round(hourlytxns(6) * W)
  Else
    tellers(6) = Round(hourlytxns(6) * W / 15 + 0.5, 0)
  If tellers(7) = Round(hourlytxns(7) * W / 15, 0) Then
    tellers(77) = Round(hourlytxns(7) * W)
  Else
    tellers(7) = Round(hourlytxns(7) * W / 15 + 0.5, 0)
  If tellers(8) = Round(hourlytxns(8) * W / 15, 0) Then
    tellers(8) = Round(hourlytxns(8) * W)
  Else
    tellers(8) = Round(hourlytxns(8) * W / 15 + 0.5, 0)
  End If
  End If
  End If
  End If
  End If
  End If
  End If
  End If
  End If
  Dim tcount As Integer
  For tcount = 0 To 8
  HourlyTellers.AddItem tellers(tcount)
  Next
  Dim totaltellers As Integer
  For tcount = 0 To 8
  totaltellers = totaltellers + tellers(tcount)
  Next
  txtTotalTellers.Text = totaltellers
  Else
  MsgBox "Please Enter an Appropriate Transactions
Value", vbCritical, "Error"
  txtTxns.Text = ""
  txtTotalTellers.Text = ""
  txtTxns.SetFocus
  End If
End Sub
Public Function Roundup(L As Double) As Long
  Dim W As Long
  W = 0.7073
  L = hourlytxns(2) * W
  If L = Round(L, 0) Then
  Roundup = hourlytxns(0)
  Else
  Roundup = Round(L + 0.5, 0)
  End If
End Function
Private Sub cmdSaturday_Click()
  frmSaturday.Show
End Sub
Private Sub cmdThursday_Click()
  frmThursday.Show
End Sub
Private Sub cmdTuesday_Click()
  frmTuesday.Show
End Sub
Private Sub cmdWednesday_Click()
  frmWednesday.Show
End Sub
```