UNIVERSITY OF NAIROBI

SCHOOL OF COMPUTING & INFORMATICS

DATA WAREHOUSING:

A Case of a Public University in Kenya

By

EMURUGAT ALFRED

P56/P/8294/2004

Supervisor

CHRISTOPHER MOTURI

May 2012

Submitted in partial fulfillment for the requirements of degree of Master of Science in Information Systems in the University of Nairobi

DECLARATION

I Alfred Emurugat do declare that this research project as presented in this report is my original work and has not been presented for any other university award.

Signed..... Date: 26/9/2012

Name: Alfred Emurugat

Reg No. P56/P/8294/2004

This research project has been submitted as part fulfillment of the requirements for the Master of Science in Information Systems of the University of Nairobi with my approval as University Supervisor.

Signed..... Date: 4 october 2012

Christopher Moturi

School of Computing and Informatics

University of Nairobi

ABSTRACT

Research has shown that only a small fraction of data in an organization is timely available to decision makers. Today a set of significant new concepts and tools have evolved into new technology that makes it possible to address the problem by providing all the key people in the organization with access to whatever level of information needed for the organization to survive and prosper in an increasingly competitive world.

The term that has come to characterize this new technology is "data warehousing". Data warehousing technology makes it easier for decision makers to analyze information in an organization without the help from Information Technology (IT) people. The main objective of this study was to determine the level of independency between top decision makers with IT personnel in accessing, analyzing and reporting data. The other objectives were to determine whether there is an easier and quicker way to analyze data, design, and implement a solution to be used to turn analysis into a report with little or no help from IT department

This study used a longitudinal research design and the mode of data collection was mainly through questionnaires, interviews and analysis of existing documents. The population considered was about 430 members of staff in the entire college. A sample size of 80 staff members was selected. Both primary and secondary data were collected. The primary data was used to establish the level of dynamic access to college data for analysis by top college decision makers, and how easily they can transform their analysis into a report. Secondary data consisted of college's reports and documents which included among others college strategic plan 2010-2013, student enrolment data ranging between January 2008 and December 2011. Data collected was analyzed using descriptive analysis, cross-tabulations, correlation analysis, and regression analysis.

The results of the survey showed that top decision makers in the college understand what data means, but they don't have easy access to the data. They depend on IT personnel to access data that they need to analyze and generate reports. As result they tend to have an uneasy relationship with IT people. However, for any serious analysis to be performed, decision makers need dynamic access to all the data and easily transform their analysis into a report without help from IT people.

The study therefore, found statistically significant to design and implement single multidimensional-cube, also known as Data Warehouse, containing student fee collection information from various courses between January 2008 and December 2011. With a data warehouse, top decision makers in the college can then dynamically access all the data they require, quickly analyze it, and easily transform their analysis into a report with little or no help from technical personnel.

Data warehouse technologies are applied within Business Intelligence (BI) systems. With the right access to data, BI software based on Excel turns ordinary spreadsheets into a flexible, powerful and inexpensive BI system. BI programs give Excel users significant power and flexibility, while eliminating many problem associated with data analysis and reporting.

ACKNOWLEDGEMENTS

I would like to pass my sincere appreciation to my project supervisor Mr Christopher Moturi who guided me throughout the project work. I would like also to extend my gratitude to the entire panelists: Dr Agnes Wausi, Mr Daniel Orwa, Mr Joseph Ogutu and Prof. Peter Waiganjo for the advice they offered on how to carry on with this project to completion.

To my family members and my close friends, thank you for being with me, no matter what the difficulty I faced. Finally I wish to thank my fellow students for the support and assistance they offered.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
TABLE OF CONTENTS iv
LIST OF TABLES v
LIST OF FIGURES vi
LIST OF ABBREVIATIONS vii
CHAPTER 1 INTRODUCTION 1
1.1 Background 1
1.2 Problem Statement 2
1.3 Objectives 3
1.4 Research Questions 3
1.5 Justification 3
1.6 Project Scope 4
1.7 Assumptions and Limitations 4
1.8 Chapter Summaries 4
CHAPTER 2 LITERATURE REVIEW 5
2.1 Area of Decision Making 5
2.2 University Data Warehousing 11
2.3 Data Warehousing 12
2.4 Ways of Implementing a Data Warehouse 15
2.5 Data Warehouse System Technologies 18
2.6 Data Modeling Techniques 19
2.7 OLAP Architecture 21
CHAPTER 3: METHODOLOGY 22
3.1 Research Design 22
3.2 Requirements 23
3.3 Design 23
CHAPTER 4: RESULTS and DISCUSSIONS 28
4.1 Implementing Data Warehouse 28
4.2 Accessing Data in a Cube (Data Warehouse Database) 33
4.3 Analyzing and Reporting data in Data Warehouse database 36
4.4 Quantitative Research Findings 43
CHAPTER 5: CONCLUSION and RECOMMENDATIONS 58
5.1 Overview 58
5.2 How Research Objectives were Achieved 58
5.3 Research Achievements 60
5.4 Summary of Research Findings 61
5.5 Recommendations 62
5.6 Limitations and Constraints 63
5.7 Areas for future work 63
5.8 Summary 63
REFERENCES i
APPENDIX iv

LIST OF TABLES

Table 2. 1: Four cognitive styles for taking decisions (Wierschem et al, 2003) 6
Table 2. 2: Stages in making a decision [Simon, 1965] 8
Table 2. 3: Structured vs Unstructured Decisions 10
Table 2. 4: Useful ways of categorizing Decisions 10
Table 2. 5: Data warehouse systems vs Operational systems 13
Table 2. 6: Data warehouse characteristics according to Inmon 16
Table 4. 1 Yearly Fee Collection 36
Table 4. 2 Fee Collection per Course per Year 37
Table 4. 3 Fee Collection per Degree Course Name per Year 38
Table 4. 4 Most Reliable and Highly Income Generating Course 39
Table 4. 5 Course Hierarchy 40
Table 4. 6 Agreement with statements made about Data Analysis and Reporting 44
Table 4. 7 Reports are produced with the help of Technician 46
Table 4. 8 I cannot dynamically Access all required data 47
Table 4. 9 Dissimilar reports produced later 48
Table 4. 10 Reports often contain errors due to source data 49
Table 4. 11 Reports are not developed quickly and easily 50
Table 4. 12 Report users don't have direct access to operational data(base) 51
Table 4. 13 Overall, how would you rate the quality of data analysis and reporting 52
Table 4. 14 Overall satisfaction level of analyzing and reporting data 53
Table 4. 15 Have you used Ms Excel 54
Table 4. 16 Have you used Excel as BI 55
Table 4. 17 Ms Excel can be BI tool 56

LIST OF FIGURES

Figure 2. 2: Information characteristics for managerial decisions 9
Figure 2. 4: *Top-Down (Enterprise Data Warehouse) approach* 17
Figure 2. 5: *Bottom-Up (Data Mart) approach* 18
Figure 2. 6: *Star Schema* 20
Figure 2. 7: *OLAP Architecture* 21
Figure 3. 1: Data Warehouse Environment 24
Figure 3. 2: Business Lines 25
Figure 3. 3: Star Schema 25
Figure 3. 4: Star Schema (Conceptual) 26
Figure 3. 5: Star Schema (Logical) 27
Figure 3. 6: Finance Fact Table Components 27
Figure 4. 1 Startup Analysis Manager 29
Figure 4. 2 Create Database 29
Figure 4. 3 Assign a Database Name 29
Figure 4. 4 Select Provider 30
Figure 4. 5 Create a cube using a wizard 30
Figure 4. 6 Cube Wizard 31
Figure 4. 7 Finish Cube Wizard 31
Figure 4. 8 Cube Editor 32
Figure 4. 9: Storage Design 32
Figure 4. 10 Process Cube 33
Figure 4. 11 Ms Analysis Server 34
Figure 4. 12 OLAP Reporting Tool for Excel 35
Figure 4. 13 Fees Collection per Year 36
Figure 4. 14 Fee Collection per course per year 37
Figure 4. 15 Revenue Collection Per Degree Course Name per Year 38
Figure 4. 16 Most Reliable and Highly Income generating Course 39
Figure 4. 17 Fee Collection per Semester per Year 41
Figure 4. 18 Fee per Level per Semester per Year 41
Figure 4. 19 Fee Collection per Month per Year 42
Figure 4. 20 Drill Down 42
Figure 4. 21 Drill Through 43
Figure 4. 22 Agreement with the statements made about data analysis and reporting 45
Figure 4. 23 How long it takes to create a report 46
Figure 4. 24 Reports are produced with the help of Technician 47
Figure 4. 25 I cannot dynamically Access all required data 48
Figure 4. 26 Dissimilar reports produced later 49
Figure 4. 27 Reports often contain errors due to source data 50
Figure 4. 28 Reports are NOT Developed Quickly and Easily 51
Figure 4. 29 Report users don't have direct access to operational data(base) 52
Figure 4. 30 Overall, how would you rate the quality of data analysis and reporting? 53
Figure 4. 31 Overall satisfaction level of analyzing and reporting data 54
Figure 4. 32 Have you used Ms Excel before 55
Figure 4. 33 Have you used Excel as BI 56
Figure 4. 34 Ms Excel can be BI tool 57

LIST OF ABBREVIATIONS

AQRT	Average Query Response Time
BI	Business Intelligence
DFM	Dimensional Fact Model
DTS	Data Transformation Services
DW	Data Warehouse
DWS	Data Warehouse System
ER	Entity Relationship
ETL	Extraction Transformation Loading
HOLAP	Hybrid OLAP
IT	Information Technology
KB	Kilobytes
MDX	Multi-Dimensional eXpressions
MOLAP	Multidimensional OLAP
MS	Microsoft
ODS	Operational Data Store
OLAP	OnLine Analytical Processing
OLTP	OnLine Transaction Processing
OWB	Oracle Warehouse Builder
ROLAP	Relational OLAP
SSNF	Snowflake Schema Normal Form
SQL	Structured Query Language

CHAPTER 1
INTRODUCTION

Majority of businesses today including educational institutions and government agencies, are realizing the need for easy and quick access to the right information (Akintola et al, 2011). Decision makers need timely and accurate information to make sound decisions. They also need to have certain mental way of processing and appreciating information. Today, a key to survival in any business world is being able to analyze, plan and react to the changing business conditions as fast as possible (Furlow, 2001)

The purpose of this chapter is to give a background of the research. The chapter shall specifically discuss the objectives, scope and justification of the project.

1.1 Background

The Government of Kenya's overall development strategy is to provide basic but quality and equitably accessible education and training to all Kenyans. One of the main aims of education is to enhance the ability of Kenyans to preserve and utilize the environment for productive gain and sustainable livelihood.

With the introduction of free primary education and subsidized secondary education in Kenya more students qualify to join public universities. About 250,000 students graduate every year from Kenyan secondary schools. Out of these, approximately 81,000 qualify to join public universities but only 24,000 are admitted for regular programmes and the remaining 56,000 absorbed through parallel programmes and admissions to private universities. This population posed steep competition for places in public universities. One of the government's responses has been to transform some colleges or polytechnics into universities in a bid to cater for the rising number of school leavers that are looking for higher education opportunities. As a result, the growth of Kenya's University education is tremendous.

Currently, there are 30 universities in Kenya, 7 of which are public and 23 private. The public universities have over 24 constituent colleges. One of the constituent colleges was used as case for this project.

The constituent college was founded in late 1940s as a training institution for corporate students from East Africa Community. Over the period, the college has undergone structural adjustments and upgrades, from a certificate college to a public university constituent college. The College has four faculties namely: Engineering, Information Science and Technology, Mass Media and Communication, and Business. The College offers various degrees, diplomas and certificates as well as short courses.

The constituent college uses both operational systems and ERPs to manage its day-to-day's operations. These systems may include among others: financial management systems; human resources management systems; student information systems; library systems; examination management systems, course management systems, etc. These applications typically run in an on-line transaction processing (OLTP) computing environment based on business functional areas.

However, the college lacks systems that will foster decision making. Although ignored, these systems are also critical for college's survival, especially in a competitive environment. The college needs systems that can unlock its data in various ERPs and other operational systems. Systems that can consolidate and analyze data easily, create ad-hoc queries, quickly generate reports in order to sustain or gain an advantage in today's competitive and information based economy.

1.2 Problem Statement

Top decision makers in the college understand what data means but they don't have easy access to the data. The current user interface does not allow them to explore, analyze data, and create reports quickly or easily without the help from IT people. Moreover, reports produced are filled with errors because source data contained errors (Akintola et al, 2011).

The ad hoc analysis and reporting impossibilities significantly hinder decision making process. This study therefore sought to establish the level of dependency between top decision makers and IT personnel in accessing, analyzing and reporting data. It also sought to determine the most valuable and efficient way to analyze data quickly and turn analysis into a report with little or no help from IT department.

1.3 Objectives

1 To determine the level of dependency between decision makers and IT professionals in the college in accessing, analyzing and reporting data

2 To establish the most valuable technology for the college users to dynamically access, analyze and report data easily and quickly

3 To design and develop a single multidimensional-cube (data warehouse) database

4 To implement an On Line Analytical Processing (OLAP) program that provides all the key people in the college with access to whatever level of information needed for the college to survive and prosper in an increasingly competitive world.

1.4 Research Questions

1. What managerial gaps are there which hinder decision making process in the constituent college?
2. What is the most valuable technology for the university users in accessing, analyzing and reporting data easily and quickly?
3. How will the technology be implemented to narrow the managerial gaps in the college?
4. How can the technology be optimized to provide timely and reliably information to the college?

1.5 Justification

Data in a data warehouse contains the properties such as subject-oriented, integrated, time-variant and non-volatile. This data is much accurate and certainly more consistent and designed in a way that it can adapt after its source data has been changed

Data warehouse contain non-volatile data which is read-only and cannot be modified by users. This data has much less chance of errors.

Data warehouse is a technology that makes it easy to analyze college information with little or no technical knowledge. With a data warehouse it is possible for business managers to find and format information for decision making without the help from IT people

1.6 Project Scope

This project was designed to implement a single multidimensional-cube (data warehouse) database in a constituent college. This cube covered student fee collection from various courses between January 2008 and December 2011. Through such cube, the project described its source data, data mapping, data profiling and data cleansing. It also illustrated how data was populated into both dimensional and fact objects, and used a simplified Excel-based Business Intelligence tool to analyze and display the results of the cube.

1.7 Assumptions and Limitations

This project assumed that the prevailing environment, existing resources and/or equipments and manpower will meet 80% of the required resources to realize data warehousing system.

The limitation of the research was that data warehouse development is so complex and involving that the time and resources available were insufficient

1.8 Chapter Summaries

This report is organized as follows:

Chapter 1 gives about the background of the research problem, objectives and the scope, justification, and limitations of the project.

Chapter 2 presents literature review of the underlying concepts, characteristics, component, technologies and modeling techniques of data warehousing within the scope of this project.

Chapter 3 presents the methodology adopted in developing a data warehouse, covering sampling, collection and analysis of data. It also covers the approach used in developing the prototype data warehouse.

Chapter 4 presents a detailed analyzes of the results of the investigation

Chapter 5 describes how research objectives were addressed, summary findings and recommendations of this investigation study and suggests areas for future work.

CHAPTER 2
LITERATURE REVIEW

This chapter outlines a brief overview of the areas of decision making in an organization. The chapter also explains the concepts, fundamentals and architectures involved in this project.

2.1 Area of Decision Making

No decision is taken in isolation. Decisions are taken by decision takers who have certain organizational objectives in mind, have certain background and have certain mental way of processing and appreciating information. Moreover, these individuals have personal interest that they may affect the decision making process. From the corporate point of view, information needs to be supplied to these decision takers in order that the decision taken will be most effective in the light of the organizational objectives (Dell'Aquila et al, 2007).

2.1.1 Cognitive Style and Background

According to (Curtis, 1995), "cognitive style" is term used in psychology that broadly describes the way that an individual's absorb information, process it and relate it to their existing knowledge, and use it to make decisions. Cognitive style and personal background act as filters to the information provided to the decision taker.
In outline, one approach to cognitive style (Ackoff, 1989) regards individuals as falling into one of the two categories in the way to absorb information.

At one extreme some people take in information best if it is highly detailed and specific, often quantitatively based. The various elements of information need not be linked as whole. The other group absorbs information in a holistic way, which is in a less concrete way, preferring general facts, suppositions and "soft data" linked as a whole.

After obtaining information the decision must be taken. Once again there appear to be two distinctive styles. One style will involve itself in high degree of analytic thought in reaching a decision. This group will be capable of providing detailed justifications often involving quantitative reasons in support of final decisions. The other group will (Wierschem et al,2003), move on intuition, experience, rules of the thumb and judgment. There will be a concentration on looking at the situation as a whole rather than parts of it independently.

This group will find it different to provide justification for recommended decisions. The combination of user information absorption and decision-taking style is shown in the table below.

	Information Absorption Style		
Decision – making style		Detailed	Holistic
	Analysis	1	2
	Intuitive	3	4

Table 2. 1: Four cognitive styles for taking decisions (Wierschem et al, 2003)

It should be noted that if information is presented in a way that is not conducive to the cognitive style of the recipient then it will not be fully utilized in a decision. In line with these, it is necessary and I should take into account the range of cognitive styles of those for whom the information is provided.

The background of the decision makers is also a powerful influence on the way information is perceived. Differing subject specifications will lead individuals to judge different aspects of the information as being more or less relevant for making decisions.

For instance, accountants will tend to concentrate on numerical information with which they are familiar. They will require the numerical information to be compiled and presented in a standard manner compatible with their expectation and training (Barry,1997). They may ignore details of the organizational structure and management styles. It is possible that they even fail to perceive the information when it is presented.

In contrast the organizational specialist may not understand the importance of numerical, financial and cost aspect of the business organization. This is quite understandable as the socialism of each only give them limited model through which they to perceive and organize information in a way that is relevant to making decision.

2.1.2 A Model for Decision Making

The process of taking a decision can be described as falling into several stages (Wierschem et al, 2003). These stages provide a framework within which decisions can be viewed. To be successfully executed each of the stages will require different types of information.

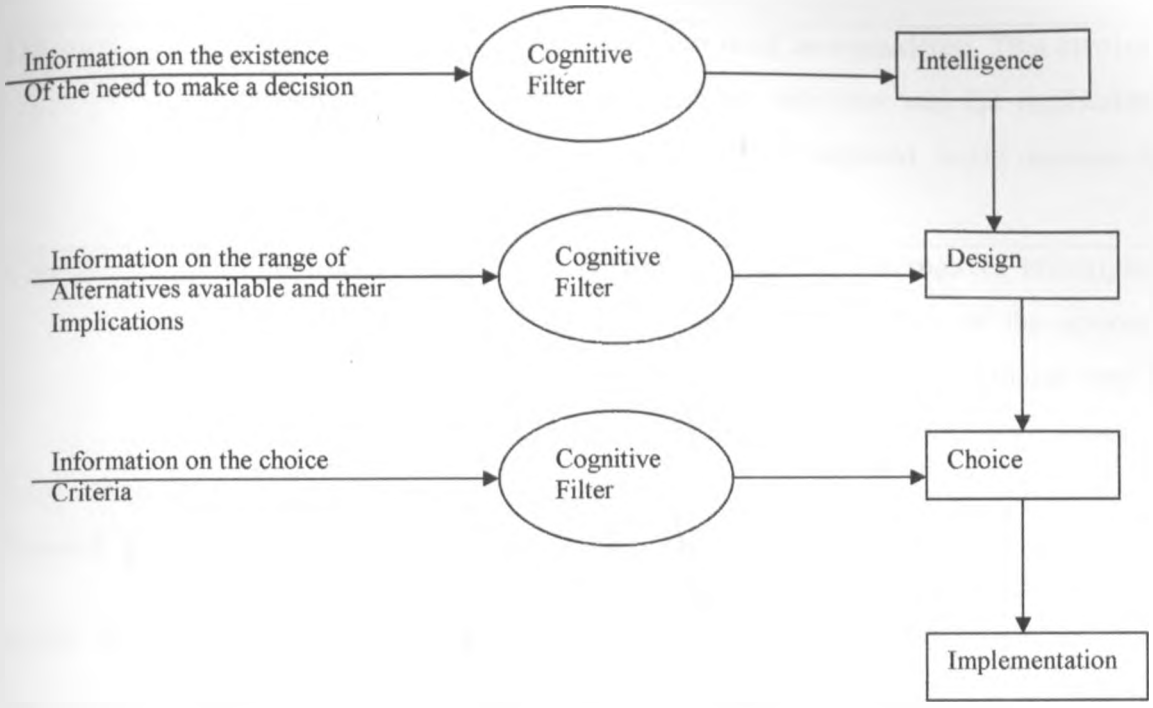


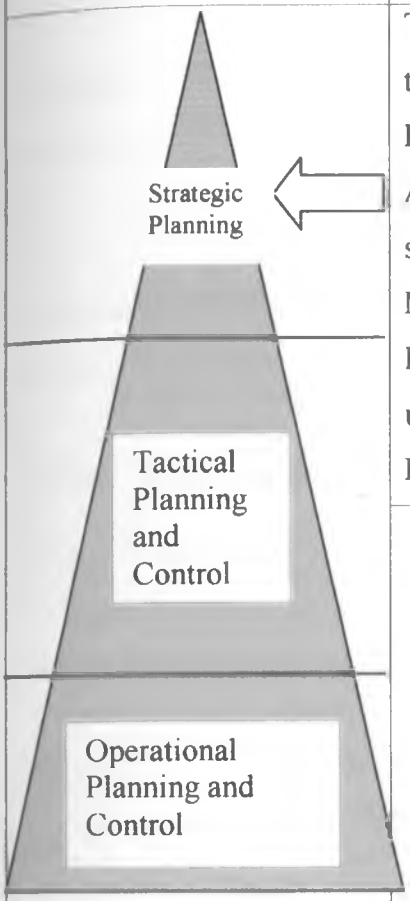
Figure 2. 1: A framework within which decisions can be viewed (Wierschem et al, 2003)

Stage	Description
Intelligence	The decision maker needs to be made aware of the existence of a problem that requires some sort of decision. Information needs to be presented in a manner conducive to this recognition
Decision	Alternative solutions to the problem must be considered. This involves the recognition of the range of acceptable decisions and the implications to each. At this stage information needs to be supplied to aid decision maker in predicting and evaluating these implications
Choice	This involves the choice between the various alternatives investigated in the previous stage. If there has been a full analysis of the options this should be a straightforward stage. Otherwise the decision maker may have to choose between incomplete and perhaps incomparable alternatives.
Implementation	The chosen decision is carried out

Table 2. 2: Stages in making a decision [Simon, 1965]

2.2.1 Levels of Managerial Decision Making

According (Gallegos et al, 1999), there are three levels of managerial activities that is strategic planning, tactical planning and control, and operational planning and control. The figure 2.3 shows the characteristics of information supplied for various levels of managerial activities. These activities need not necessarily be carried out by different people



	Information Characteristics	Involves
	Time horizon: Long term Level of details: Aggregated summarized Source: Mainly External Degree of certainty: uncertainty Frequencies: Infrequent	Decisions on what markets to move to Whether to diversify production Determination of organizational objectives How to structure the company's finances Allocation of resource should to major functions
		Allocation of resources within department budget Decision on medium term work scheduling & forecasting Planning medium term cash flow Monitoring of actual production and expenditure Analysis of variances and actions taken
	Time horizon: Immediate Level of details: Highly detailed Source: Interior Degree of certainty: certainty Frequencies: frequent	Hiring and firing Control of inventory and production levels Pricing decision Aspects of credit control over customers

Figure 2. 2: Information characteristics for managerial decisions

2.1.3 The Structure of Decisions

(Simon, 1965) makes a simple yet important distinction between structured (routine) and unstructured (non-routine) decisions.

The table 2.3 shows the difference between the two structures of decisions in relation to procedures to be followed and the information required in making decision

	Structured decisions	Unstructured decisions
Decision procedures	Decisions are governed by clear rules. The decision procedures can be expressed as a set of steps to be followed	No set of procedures or rules for arriving at the decision. The use of rules of thumb (heuristics) and 'experience'
Information required	Clearly specifiable, unambiguous and once it is obtained, the process	Unclear and assessed in relation to the decision objectives

Table 2. 3: Structured vs Unstructured Decisions

(Gory & scott-morton, 1971) have developed the ideas of (Anthony, 1965) and (Simon, 1977) to provide a useful way of categorizing decisions by comparing managerial activities against the degree of structure in a decision as shown below

Decision Classification	Strategic Planning	Management control	Operation control
Structured (Repetitive and routine)	Financial structure planning	Allocating budgets	Stock reorder decisions
Unstructured (non-routine decisions	Company reorganization	Personal management	Dealing with customer inquiries
Semi-structured (includes some characteristics of both)	Introduction of products	Analysis of performance	Short term production scheduling

Table 2. 4: Useful ways of categorizing Decisions

The degree of structure corresponds in the extent to which each of the decision making stages is structured or unstructured. A decision that is highly structured at the stage of intelligence, design and choice would count as a highly structured decision. Lack of structure during each of these three stages would mean the decision was regarded as highly unstructured. Many

decisions lie between these two extremes, being structured in some stages and unstructured in others, these are termed semi-structured.

(Simon, 1977), As the type of decision will determine the characteristic of the information that is required to make it, the analysis by structure provides a useful guide for the development of management information systems. In general the more highly structured the decisions the more likely it is that a computer system can provide useful information. In cases where the intelligence, decision and choice elements are structured the computer systems may not be only be used to provide information but also to automate the decision itself. In other cases varying degrees of decision support can be given.

2.2 University Data Warehousing

Like other businesses, a university need new approaches of providing a centralized source of information accessible across different academic units to quickly analyze problems and get satisfactory solutions, supplying the data necessary for developing the institution's strategic plan and enabling administrator to make better business decisions based on historical data available in the legacy databases (Wierschem et al 2003 , Dell'Aquila et al).

For institutions to gain competitive advantage over others and to help make better decisions, data warehousing cum data mining are now playing significant role in strategic decision making. In a case study of Federal University of technology Student-Course Management systems (Akinttola et al 2001), it demonstrated the process of data warehouse and data mining applications using SQL server business intelligence development tools in academic environment. It is important to note at this time that the same technique can be applied in other organizations wishing to implement business intelligence as part of their strategic decision support operations.

The proponents (Akinttola et al 2001) argue that organizations, particularly in Nigeria began to implement the product of their case study as part of their strategic decision making process tool.

McMaster University (Gallegos et al 1999), data warehouse was built to support the University's mission statement, hence, needs to be flexible and innovative Information Technology the will ensure the transfer and sharing of knowledge to inform and enhance all

levels of decision making at the institution. The Data Warehouse was deployed incrementally along with end-user, web-based query and reporting tools using the SAS Intelligence Value Chain. Like many other organizations, McMaster University's transactional applications do not store data in easily accessible data models that can be easily transformed into comprehensive, meaningful information to support evidenced-based decision-making. According (Lechtenbörger & Vossen, 2003), an external consultant was engaged to design a conceptual data model that included all current and anticipated information needs across the enterprise. This was done after educating the stakeholders about data warehousing best practices and guiding them on developing a detailed requirements analysis.

From the implementation of McMaster University data warehouse, it is important to echo the use of consultative approach where the stakeholders took the central role leading to a product designed to answer business questions posed by business users. This approach not only helped in knowledge transfer through training and mentoring, but also in assessing information needs, building data hierarchies, understanding subject matter, and in prioritizing tasks to best support decision and inform senior management.

2.3 Data Warehousing

Data warehouse, is a technology that makes it easy to analyze corporate information without technical knowledge

With data warehousing it is possible for business managers to find and format information for decision making without help from IT people. The ad hoc analysis and reporting possibilities speed up decision making significantly and therefore make organizations much more agile. However, data warehousing technologies need to be implemented carefully for business users to be able to adopt them (Pendse, 2005).

2.3.1 What is Data Warehousing

A process of transforming data into information and making it available to users in a timely enough manner to make a difference

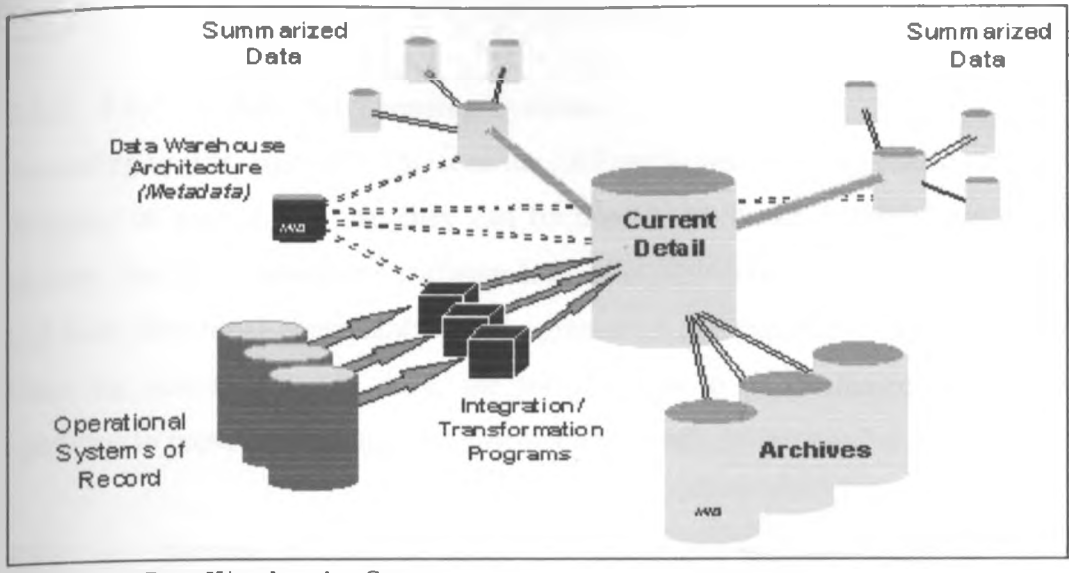


Figure 2. 3: Data Warehousing System

2.3.2 Operational Systems vs Data Warehousing Systems

	Data Warehouse Systems	Operating Systems
business point of view	Allow a company to identify opportunities for increasing revenues, and business growth	Allows a company to manage its operations at reduced cost
Size and content:	The historical information assist the company to in better understanding of its business growth in size	The operational database aims at staying small
Performance:	The true goal of data warehouse is to provide better information, or business intelligence.	In an operational environment, speed is of the essence.
Content focus:	A data warehouse focuses on cross-functional subject areas.	Focuses on small work areas, not the entire enterprise;
Tools:	A data warehouse is the land of user tools. Various tools are available to support the types of data requests discussed earlier.	Operational systems are typically structured, offering only a few ways to enter or access the data that they manage, and lack a large amount of tools accessibility for users.

Table 2. 5: Data warehouse systems vs Operational systems

2.3.3 ERPs vs Data Warehousing Systems

According to (Prasad 2006), "most of the ERP packages are very generic packages. They are designed in such a way that they can be used in an organization of the any type in any country. So, to customize the software for one organization in one country is a gigantic task and both functional consultants and programmers are required to work on ERP packages. Once the customization is done, the software has to be maintained to ensure its proper operation to provide MIS support to the management".

ERPs have become the way of doing accounting, billing, collections, marketing, strategic management and everything else essential to a company or firm. These systems brought the promise of a single, integrated processing environment to an otherwise chaotic mix of business applications tied together by a maze of complex interface programs. The promise of the ERP system was that one system would solve all the information needs of an organization. The results have been mixed, although ERP systems are good at capturing and storing information, the system lacks capabilities to analyse and report data. The ERP system is the backbone of the data processing for many organizations but business users are frustrated with the inability to get the information out of the business.

Theoretically, (Velicanu & Matei, 2007) all the data is in ERP system. Yet it remains locked, not easily retrieved for a analysis. In addition, the ERP system cannot accommodate data from other applications and external sources necessary for analytic processing of the intelligent enterprise.

Corporate data is the lifeblood of the competitive organization; ERP is the heart, it is the data warehouse that is the brain consolidating the ERP data along with the external sources of data in the creation of information and support of analysis. It is the business Intelligence environment that can unlock the information collected by the ERP and other operational systems.

ERP vendors are focusing on their own data warehousing solutions, and third-party data warehouse vendors are rallying around the ERP marketplace. The number of options is overwhelming. Matching the right tools to the business requirements and the operational landscape will ensure success.

It is clear that as the volume of data in the database increases, the performance of the database and the related applications decreases. "It is evident that we should separate the operational data from the non-operational data. Here the term archive data is not used, because if the non-operational data is archived, there is little or number use of it. But this data is very valuable resource and is too precious to be kept in some archive. It is in this situation that data warehouse comes in handy. Separating it from data in operational systems can access the primary concept of a data warehousing as that the data stored in business analysis most effectively. The most important reason for separating data for business analysis, from operational data, has always been the potential performance since the analysis processes much more complex and sophisticated. In addition to producing standard reports today's data warehousing systems support very sophisticated online analysis, including multi-dimensional analysis.

2.4 Ways of Implementing a Data Warehouse

It should be noted that there are wide varieties of different approaches to build a data warehouse. This project, nonetheless, documents four main approaches that is top-down (Enterprise Data Warehouse); bottom-up (Data Marts); hybrid and federated. The first two represent the most common strategies used by many companies to build their data warehouse systems (Weisensee, et al).

The next few sections critically examines these approaches, where there pros and cons are discussed including the differences and similarities that exists among these architectures.

2.4.1 Top-Down Approach

According to Inmon (one of the earliest and most influential practitioners) school of thought, top-down approach is based on the construction of an enterprise data warehouse. In his approach (Inmon 1996), companies first build an enterprise data warehouse in an interactive manner, business area by business area, and the underlying dependent data marts and data structures are subsequently created based on the content contained in the enterprise data warehouse.

Inmon defines a data warehouse as

> "*A data warehouse is subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process*"

Characteristics	Explanation
Subject-oriented	The data addresses a specific subject such as enrolment, examination, finance
Integrated data	The data is obtained from a variety of sources
Time-variant	The data is stored in such a way that when some data is changed, when that the data has been changed is also stored.
Non-Volatile	The data is never removed, i.e. historical data is also kept.

Table 2. 6: ***Data warehouse characteristics according to Inmon***

In a nutshell, this approach, must respond to the requirements of all the users in the organization, not for only a certain group. It views the data warehouse as the single organizational repository of enterprise wide data across all subject areas. The data warehouse serves as the '**single source of truth**' for all fact and dimension data within that organization.

From this central location of data, the data is summarized, dimensionalized and distributed to one or more dependent data marts as shown in figure 2.4. The single repository of all meta-data and meta-data rules provide less complex system. Also, the proponents of this approach believe that by delivering an integrated architecture throughout the enterprise, every subset data mart that is created from the data warehouse inherits the integrated architecture, easing maintenance for IT, allowing greater flexibility of the end-users.

Nevertheless, this approach is generally slow in its implementation because of its iterative nature. Furthermore, to achieve the 'single version of the truth', the design and development of the data warehouse needs the involvement from all functional areas within an organization.

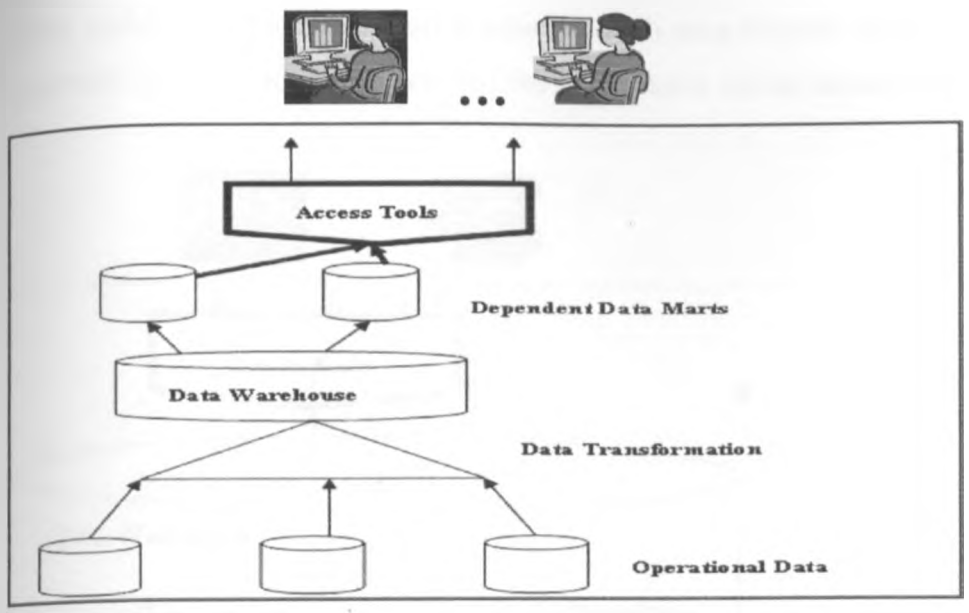


Figure 2. 4: ***Top-Down (Enterprise Data Warehouse) approach***

Beside, the top-down architecture has higher risks and levels of failure because they require a large up-front development effort and implementation of multiple, large, normalized target databases. This approach, also, demands for it to be used the technology and economic problems should be well known (Lechtenbörger & Vossen et al, 2003).

2.4.2 Bottom-Up Approach

The bottom-up approach advocated by Ralph Kimball, builds incrementally data marts towards the goal of an enterprise data warehouse.

According to (Kimball, 1996), data mart is defined as follows:

"A copy of transactional data specifically structured for query and analysis"

From the definition, data warehouse provides access to corporate or organizational data; data in data warehouse is consistent; data warehouse is not just data, but also a set of tools to query, analyze and present information; data warehouse is the place where we publish used data.

Kimball argues that, independent data marts are created with the view of integrating them into an enterprise data warehouse at some time in the future as shown in figure 2.5. The proponents of this approach maintain that, each data mart builds on the next, reusing dimensions and facts so users can query across data marts if desired, to obtain a 'single version of truth' as well as both summary and detailed data hence users do not have to "dill-through" from a data mart to another structure to obtain detailed data. They also argue that,

since architecture for a data mart is typically built on a focused scope, production status can generally be achieved in less than half the time it takes for the top-down approach.

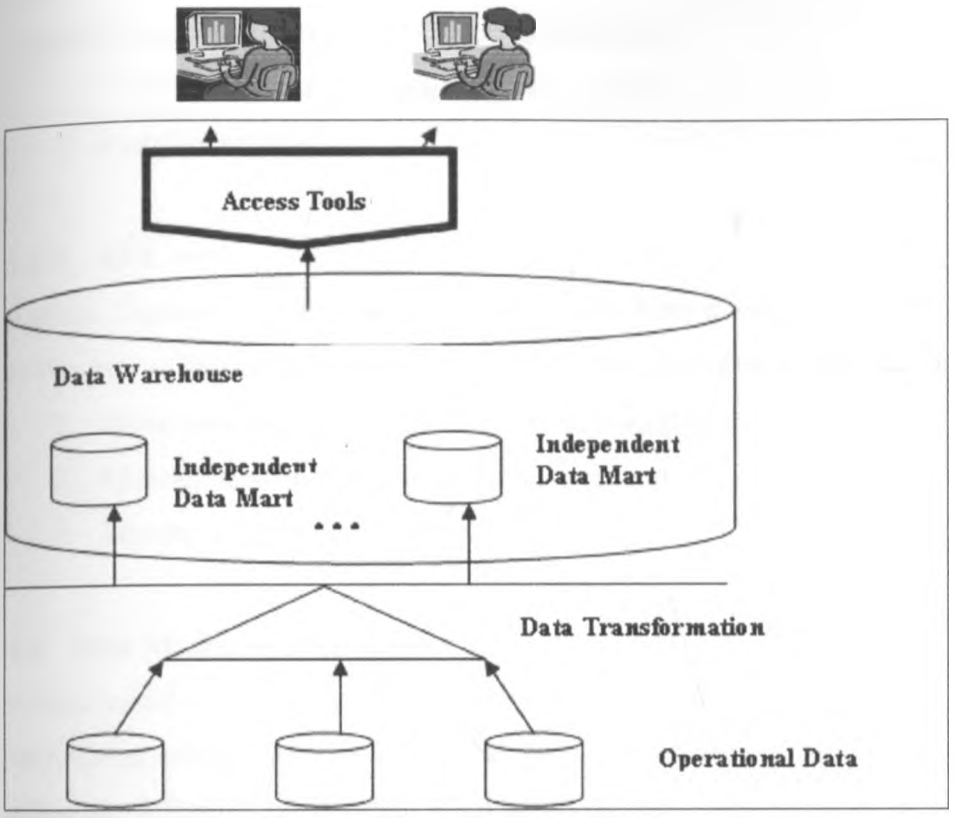


Figure 2. 5: ***Bottom-Up (Data Mart) approach***

2.5 Data Warehouse System Technologies

The following data warehouse technologies are available to support data warehouse components

2.5.1 Relational databases

These are the underlying engines that provide power to the data warehouse. Data warehouses store, manage, and manipulate huge quantities of data, often as much as hundreds of millions of rows of historical information. The relational database must provide rapid data transfer and updating, flexible and efficient indexing, the sophisticated and effective query capabilities to organize and retrieve data warehouse data.

This technology is used in data warehouse systems to:

1. Contain and manage the large quantities of data in the data warehouse database
2. Stage, clean, and transform incoming data in the data staging area.

2.5.2 Online Analytical Processing (OLAP) Databases

This technology provides an alternative to relational database technology, offering flexible data viewing, analysis and navigation coupled with fast data retrieval.

The two main characteristics of OLAP technologies:

1. An intuitive, multidimensional view of data
2. Fast data retrieval

2.5.3 ETL tools

Extract, Transform, and Load tools extract data from a data source and apply transformation to the data so that data is in a format to load into a destination data source. Tools that:

1. Move data from the source systems to a staging area.
2. Moving data from a staging area to a data mart.
3. Moving data between data marts.

2.6 Data Modeling Techniques

A data model is a combination of constructs used to organize data. Each data model provides structuring mechanisms, similar to the type constructors of programming languages, which allow the definition of new data types based on constructors applied to predefined elementary types. There are two data modeling techniques that are relevant in a data warehouse environment.

Dimensional Modeling

This model is an imperative technique of modeling the needs of a business user especially the data this user requires to accomplish business objectives. This technique can be represented as a star, and hence this model is also referred to as star schema.

Star Schema

To facilitate data retrieval and analysis, a data warehouse physically organizes data into easy-to-use structures called star schema. A data warehouse that is composed of subject-oriented data marts uses one or more schemas to represent specific business events or processes.

For example, a data mart uses a star schema to store summarized data.

A star schema contains a central fact table surrounded by many dimensions tables

Fact Table:

Used to describe a specific event in a business, such as an individual customer purpose or it can contain summarized data, such as enrollment that are summarized from student admissions transactions.

Fact table grain

The fact table grain defines the lowest level of detail stored in each of the dimensions associated with the fact table

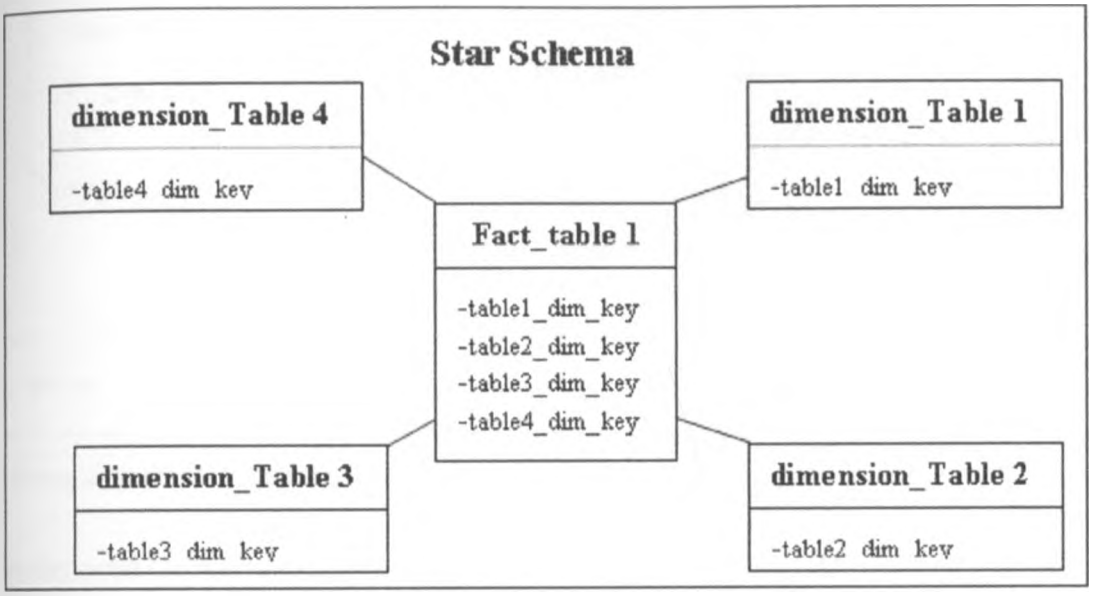


Figure 2. 6: *Star Schema*

Dimension Tables

A dimension table represents a business entity, providing the context for numeric data in the fact table. The design of dimension tables targets the analytical needs of the business user by presenting usage, descriptive information that is easy for users to browse.

2.7 OLAP Architecture

This OLAP stores the Data in relational format, while presenting it in Multi-dimensional format

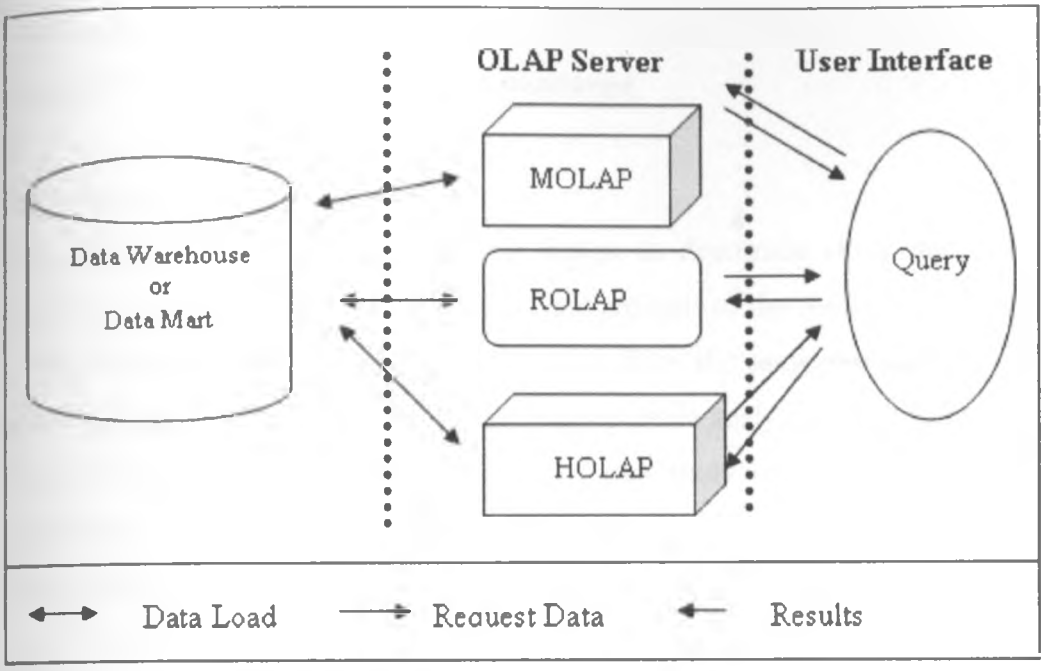


Figure 2. 7: ***OLAP Architecture***

Relational OLAP (ROLAP)

This architecture provides a multi-dimensional view of the data to the user, but stores the data in the relational database format.

Multidimensional OLAP (MOLAP)

This architecture fulfils the requirements for an analytic application, where you require to access only summarized level of data

Hybrid OLAP (HOLAP)

Hybrid OLAP (HOLAP) keeps the data for your cube or partition in the existing relational data store and to keep your aggregations in a multidimensional structure

CHAPTER 3:
METHODOLOGY

This chapter covers sampling, collection and analysis of data. It also covers the approach used in developing the proto type data warehouse.

3.1 Research Design

This study used a longitudinal survey design to determine study questions, data that is relevant to the study, data to be collected and analysis of the results. A longitudinal survey is a correlation research study that collects data from the same sample variables over a long period of time (Curtis, 1995). They have considerable analytical advantages over one-time or cross-sectional surveys. Given that part of this study soughed to establish the level of dependency between top decision makers and IT personnel in accessing, analyzing and reporting data, the longitudinal survey was the best suited for the study.

Population and Sample

The population considered by this study was the entire university members of staff. A sample size of 74 staff members was selected. Sample size was a function of homogeneity or heterogeneity of the population, and gave a clear representation of the university sub-strata decision makers

Research Data

Both primary and secondary data were collected. The primary data was used to establish institutional top management level of dynamic access to the data for analysis and how easily they can transform their analysis into a report. The student enrolment data for a period of 4 years January 2008 – December 2011 was obtain for quantitative analysis to study the trends and the performance to the institution. Other institution's reports and documents were also studied.

Method of Data Collection

The study involved the use of both qualitative and quantitative research techniques in collecting Data. These included desk study which involved identifying and analysis of documents, in-depth interviews with key respondents who included senior stakeholders from

various departments and/or sections, and use of questionnaires to ascertain the level of understanding and awareness in the stakeholders in use of Excel and data analysis and reporting concepts

Data Analysis

The data collected from survey was entered into the Statistical Package for Social Sciences (SPSS) version 10 and analyzed using descriptive statistics, cross-tabulations, correlation analysis, and regression analysis.

3.2 Requirements

University Business Requirements included

1 Tracking of fee payment by course, program, department, student, and calendar time

2 Enabling users to access, navigate, and analyze multidimensional data in an easy, natural way

3 Running top-down analysis on courses offered and students and eventually to drill down to the transaction level in the relational database

4 Predicting what types of course(s) will attract more students

3.3 Design

College Data Warehouse Data Source

The main source of data for data warehouse under consideration is from both operational systems and ERPs that the college currently uses manage its day-to-day's operations. These systems included among others: financial management systems; human resources management systems; student information systems; library systems; examination management systems, course management systems, etc. The operational data to and from these are kept in disparate operational databases.

College Data Warehouse Architecture

The data warehouse must be optimized according to the users' demands.

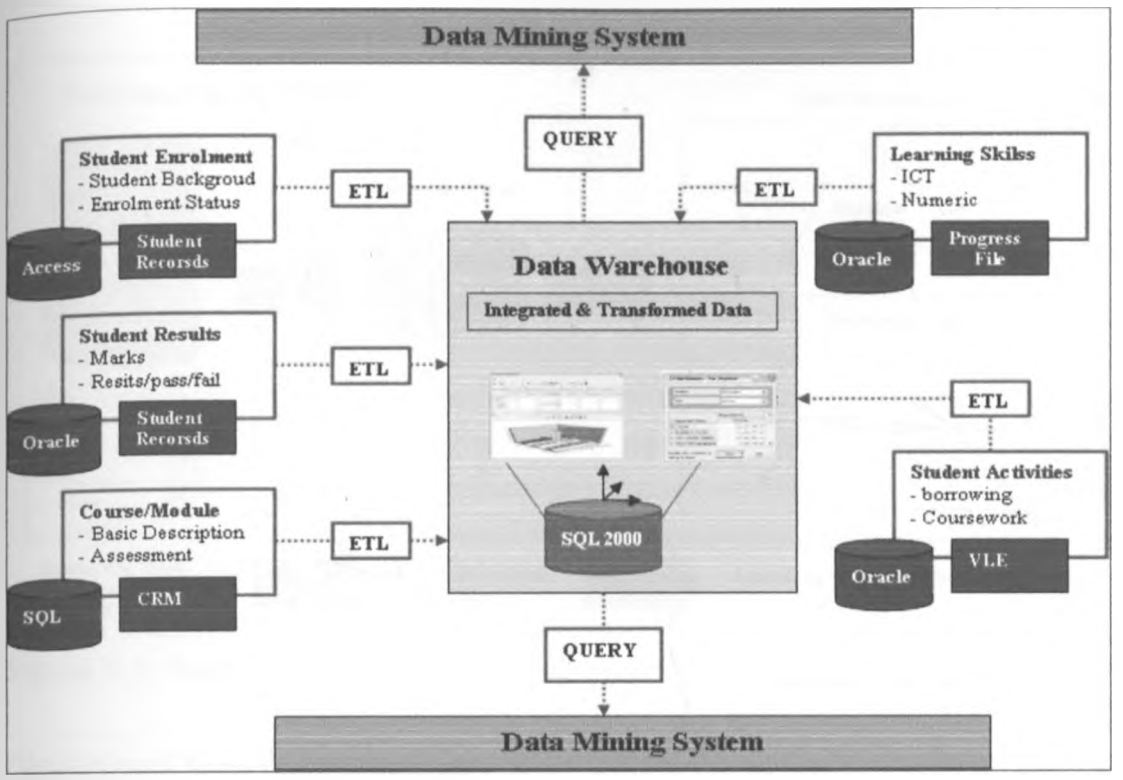


Figure 3. 1: Data Warehouse Environment

Data warehousing technology consolidates essential historical data from the above underlying disparate databases. Data stored in a data warehouse is typically extracted from operational systems and possibly from data sources external to the institution with batch processes during off-peak operational hours. Data that has been extracted is filtered to eliminate any unwanted data items and transformed to meet the data quality, security and usability requirements. It is then loaded into the appropriate data warehouse databases for access by end users who are charged with the responsibility of making informed decisions.

College Data Warehouse Subject Areas

The project found statistically significant to integrate and consolidate data into a multidimensional (OLAP) database. The OLAP database contained a single data mart – Finance data mart that tracked revenue collected from student in various courses.

The extracted data was transformed into a format that can be loaded into the destination star schemas. The staging area was used to facilitate the transformation process. The data that was already-to-load format was loaded into dimension table and fact table in each destination star

schema of the data warehouse. The transformed data was first loaded into dimension tables and then load the fact table that references the dimension.

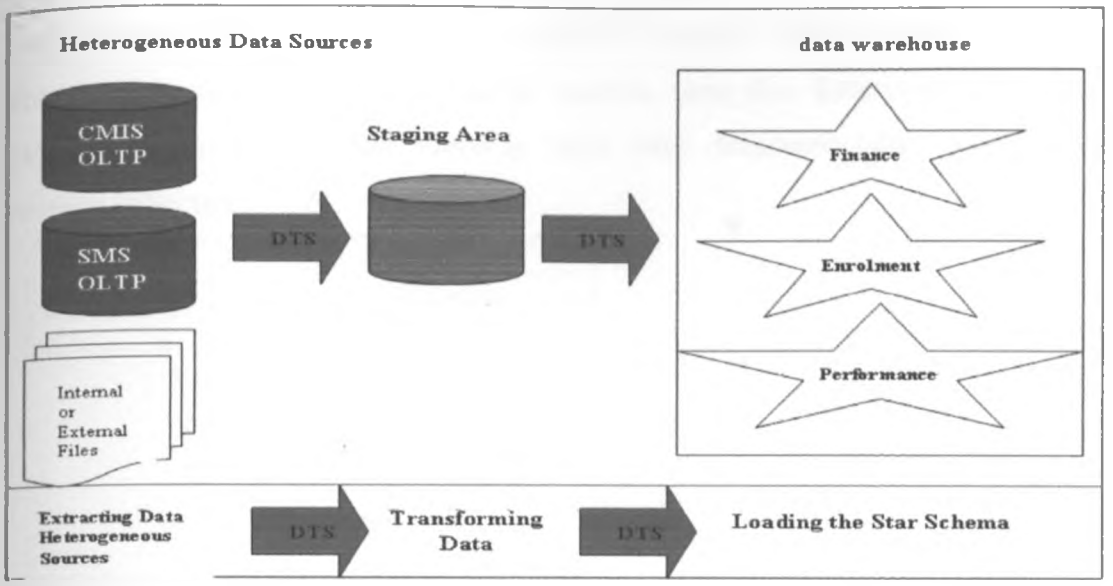


Figure 3. 2: Business Lines

Dimensional Modeling Techniques

Star Schema

To facilitate data retrieval and analysis, a data warehouse physically organized data into easy-to-use structures called star schema. Data warehouse composed of subject-oriented data mart using one or more schemas to represent specific business events or processes.

A star schema contained a central fact table surrounded by many dimensions tables

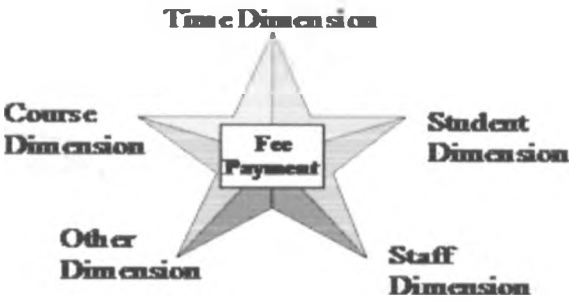


Figure 3. 3: Star Schema

The fact table contained numeric metrics at the intersection of many business entities

The above finance fact table contained enrolment data summarized by student, course and time. The dimension tables provided context for data in the fact table.

Star Schema: Conceptual Model

The dimension tables in Finance star schema are the student_dim, course_dim and time_dim tables. The student_dim dimension contained detailed student information, course_dim dimension contained courses offered to students, time_dim dimension contains the of the calendar semesters. The data stored in these three dimension tables provided context of revenue collected.

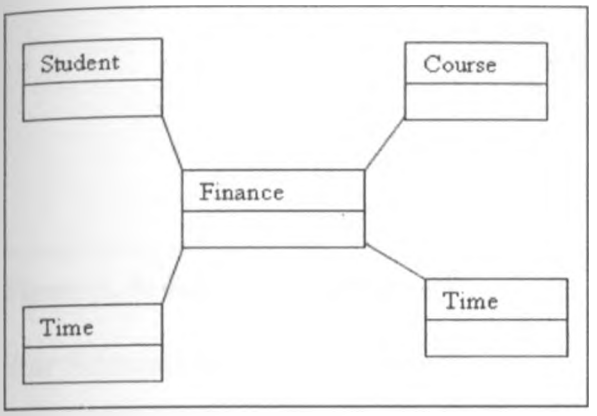


Figure 3. 4: Star Schema (Conceptual)

Star Schema: Logical Design

The design of the dimension tables targeted analytical needs of the business users by presenting usable, descriptive information that is easy for users to browse.

The figure below shows three dimension tables: student_dim, course_dim and time_dim. These dimensions represented business entities with the following lowest level of details student_id, course_id, and full_date for each dimension. Relevant attribute columns were defined and are highly correlated and descriptive.

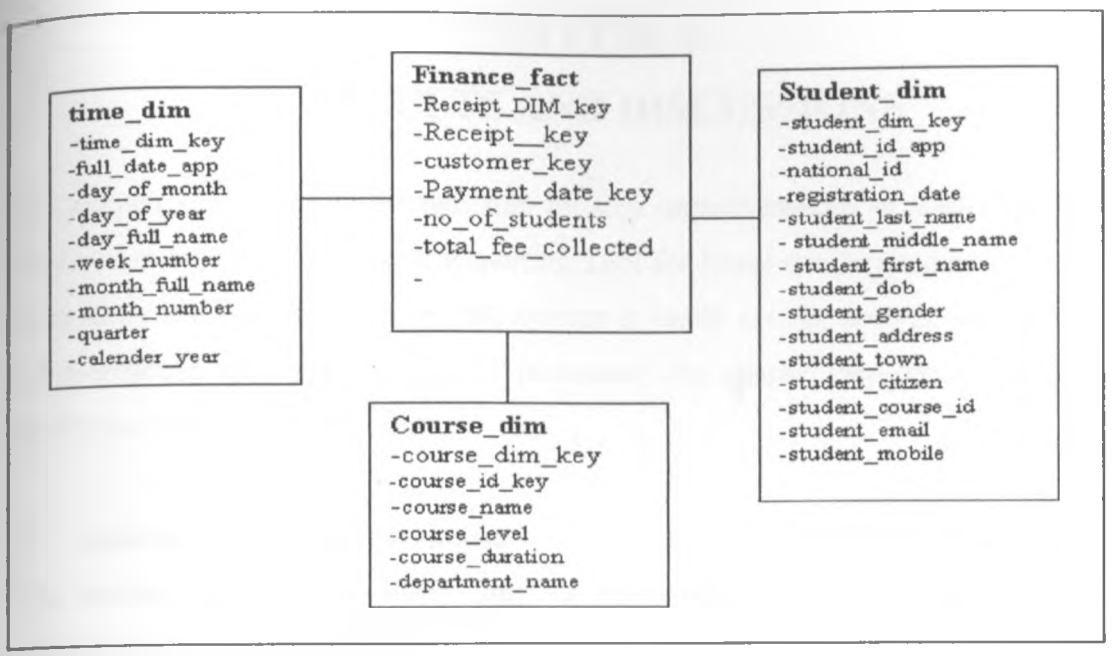


Figure 3. 5: Star Schema (Logical)

Star Schema: Physical Design

Fact Table Components

The fact table was principally made up of measures and foreign keys as shown below.

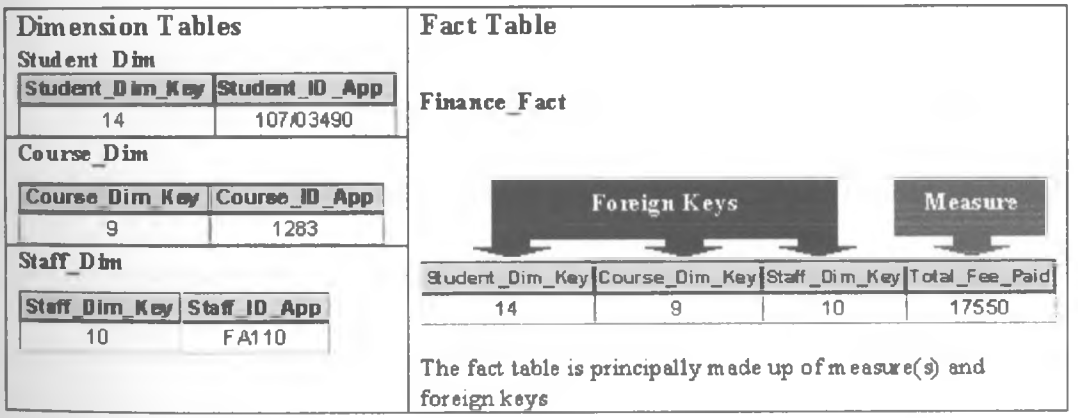


Figure 3. 6: ***Finance Fact Table Components***

Program Design

Data Transformation Services (DTS) was used to access relational and non-relational data sources to load data into the data warehouse. CMIS OLTP system resided in Ms SQL server, SMS database that resided in Ms Access 2000, data files, external files and Ms Excel spreadsheet etc as shown below.

CHAPTER 4:
RESULTS AND DISCUSSIONS

This chapter steers readers through step by step implementation of a data warehouse. The chapter also explains how OLAP Reporting Tool for Excel can be used as an example of an outstanding Business Intelligence (BI) system to easily access, analyze and report data in a data warehouse without the help of IT personnel. The chapter then concludes by discussing quantitative analysis of the survey.

4.1 Implementing Data Warehouse

This section demonstrates every step of the implementation in populating a single multidimensional (data warehouse) database, also called cube. The cube contains data, data structures and business logic. The data in a cube is stored as measures and categorized into dimensions with or without hierarchies.

In line with this project

1. The name of the cube is Fee Payment
2. Fee payment cube contained only one measure called total Amount - fee collection from students taking different courses between January, 2008 and December, 2011.
3. Dimensions contained in the cube were student, course, staff and time.
4. Time dimension consisted of a hierarchy by Year, Semester, and Month.

To populate a cube Database Transformation Services (DTS) tool was used in Microsoft SQL 2000 as shown below.

1. Start Analysis Manager

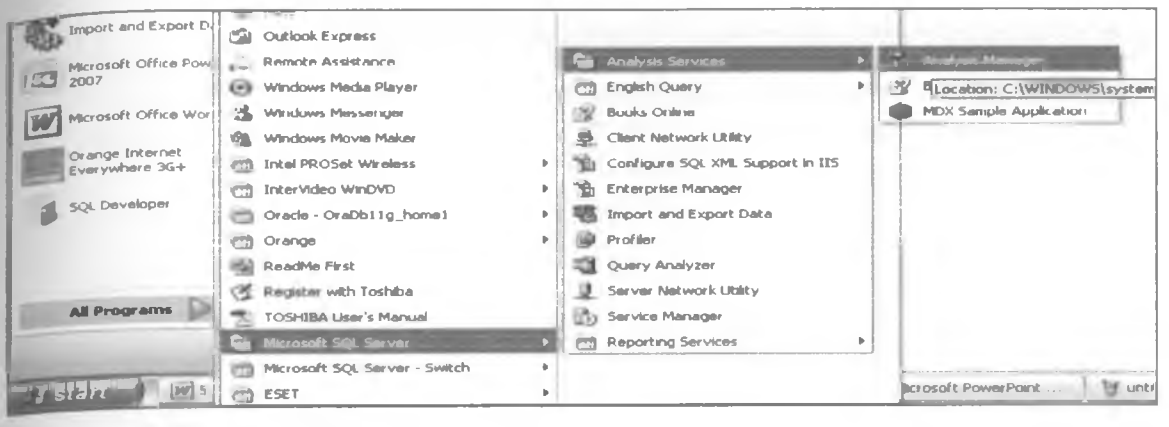


Figure 4. 1 Startup Analysis Manager

2. In the Analysis Manager window, double click Analysis Services
3. Right click local server name and select New Database

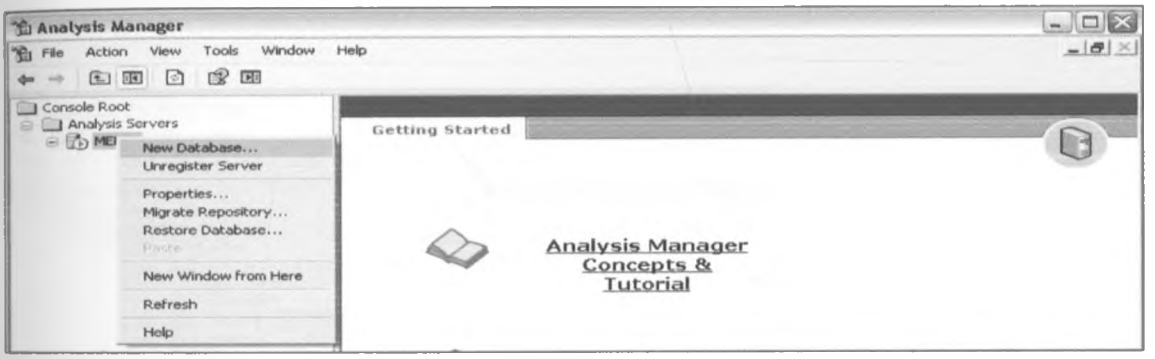


Figure 4. 2 Create Database

4. Enter name of the database you want to create e.g DW and click Ok

Figure 4. 3 Assign a Database Name

5. A database will be created. Create a connection to by right clicking data source and selecting new data source. Select the provider.

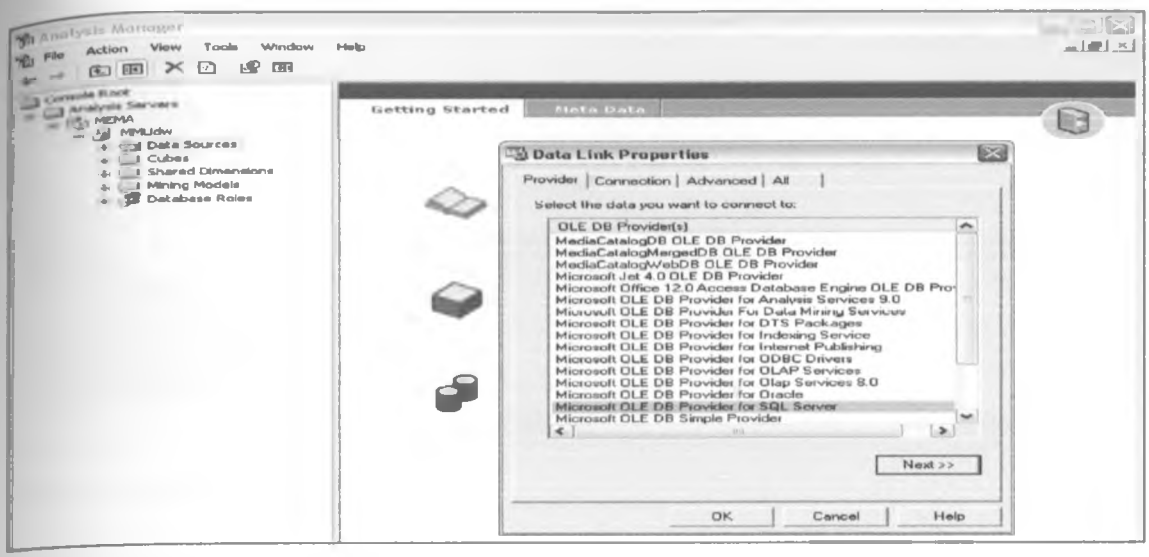


Figure 4. 4 Select Provider

6. Click Next to set the connection path and then click Ok
7. To create a cube, right click the cubes, select New Cube then wizard

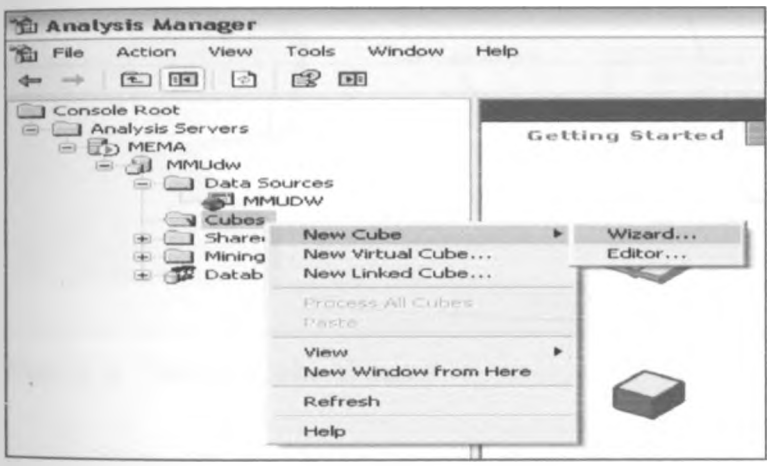


Figure 4. 5 Create a cube using a wizard

8. The Cube Wizard appears. Complete the process by specifying what is in the subsequent steps of the wizard

Figure 4. 6 Cube Wizard

9. Continue with the process until you get to the Finish the Cube Wizard
10. Enter the cube name and click finish button.

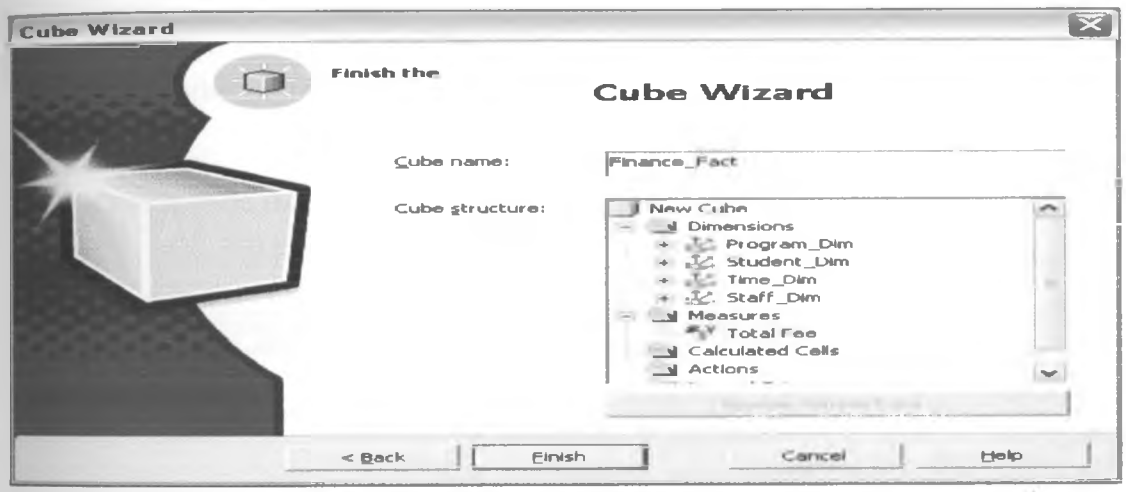


Figure 4. 7 Finish Cube Wizard

11. Cube Editor Window opens as shown below

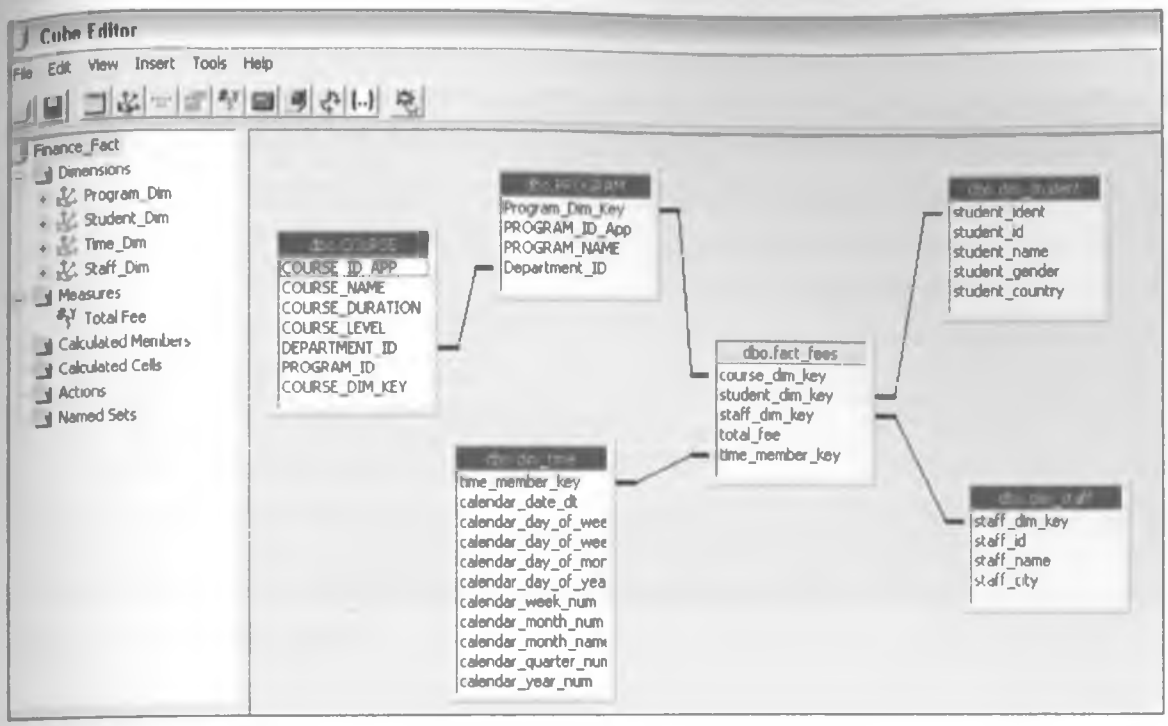


Figure 4. 8 Cube Editor

Storage Mode

Storage mode to be used was chosen and designated the amount of precalculated values to be stored. MOLAP storage mode was chosen

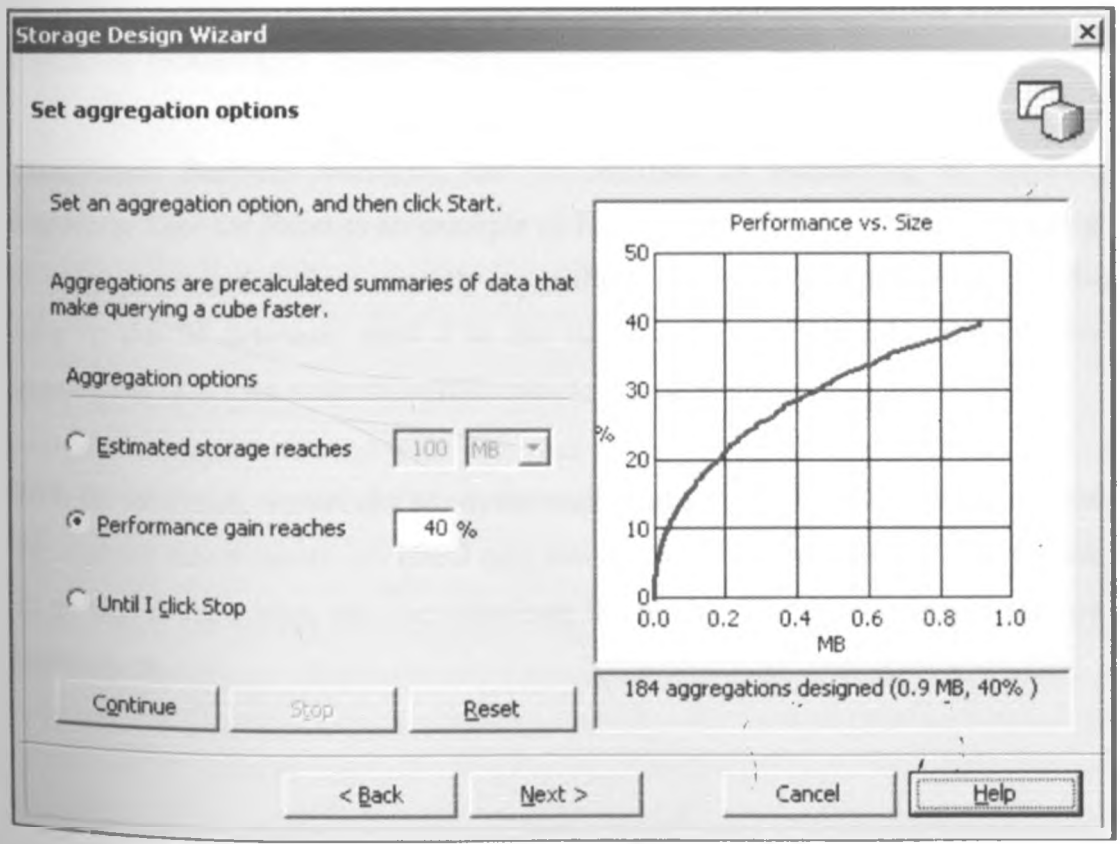


Figure 4. 9: Storage Design

12. Process the cube

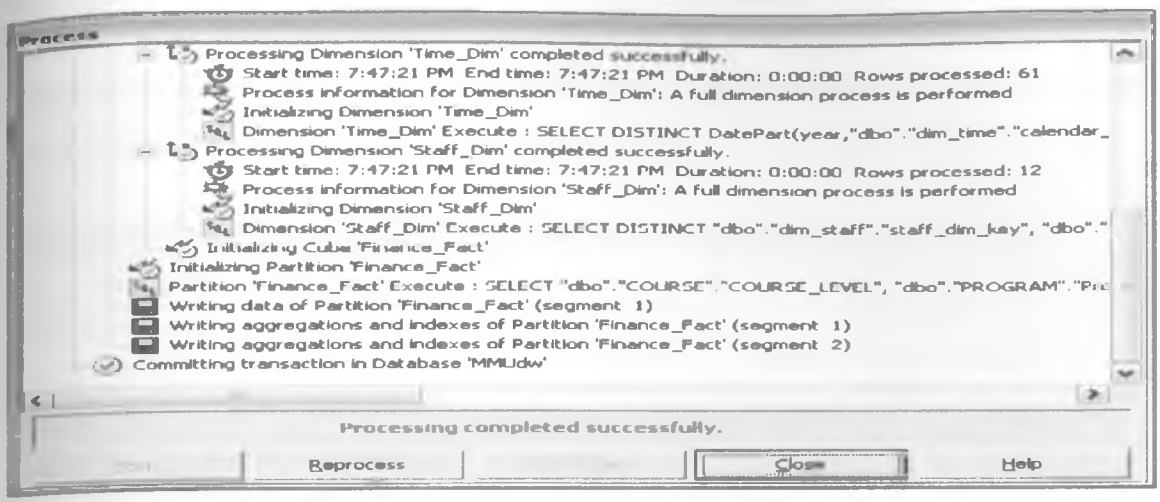


Figure 4. 10 Process Cube

4.2 Accessing Data in a Cube (Data Warehouse Database)

4.2.1 OLAP Reporting Tool for Excel

All in all, cubes are the basis for an easy, but very rich analysis and reporting experience for knowledge workers. In practice however, many BI systems, even the ones utilizing data warehouse technologies struggle with a low user adoption.

Excel-based Business Intelligent can be deployed as outstanding BI software. OLAP Reporting Tool for Excel is an example of Excel-Based BI system. This tool can give users an easy access to data from virtually any sources. An Excel spreadsheet can pull data from a cube in the BI database; send it to the add-in; receive the results; process them in the spreadsheet, and then write the results back to the BI database

With BI for Excel, workbooks are dynamically linked to BI data. Better yet, it's easy to turn the analysis into a report. An Excel user merely needs to error-check his workbook, clean it up so that it looks nice, and then distribute the results. They never need to talk with the IT Department

4.2.2 Accessing Data Warehouse Database

This tool is used to connect to Ms Analysis Services Server by specifying server name. Enter the name of the server in pleas give the name of the Ms Analysis Services Server you wish to connect to box and then click Get database button as shown below. The name of the cube(s) created appears. Select the cube name and click OK

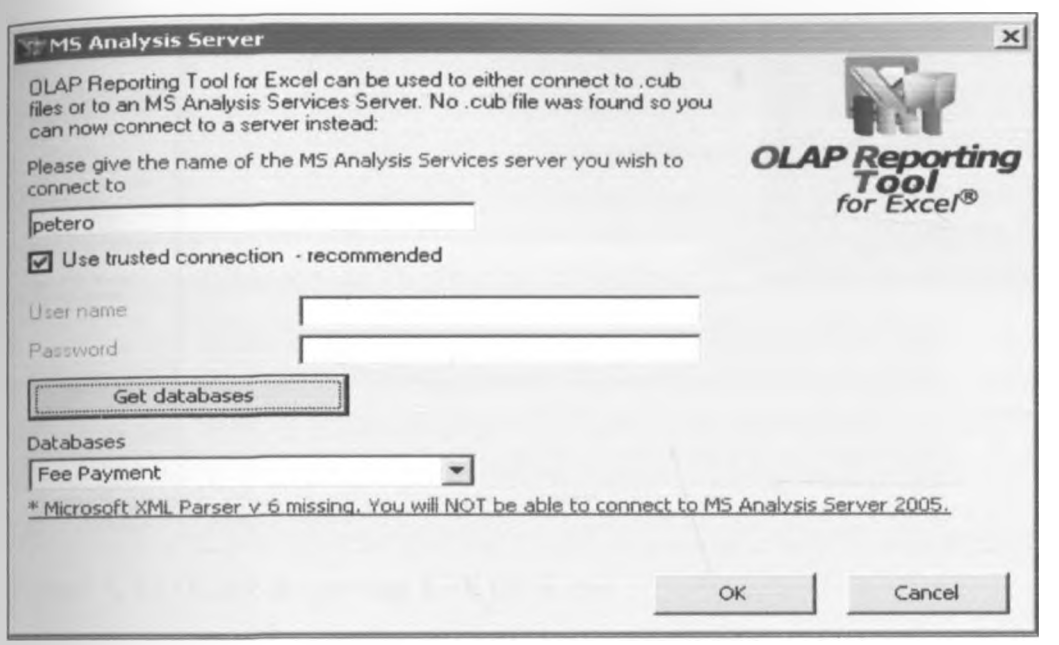


Figure 4. 11 Ms Analysis Server

4.2.3 OLAP Reporting Tool for Excel working Environment

Excel-friendly OLAP programs give Excel users significant power and flexibility, while eliminating many problems associated with writing values to cells as shown below

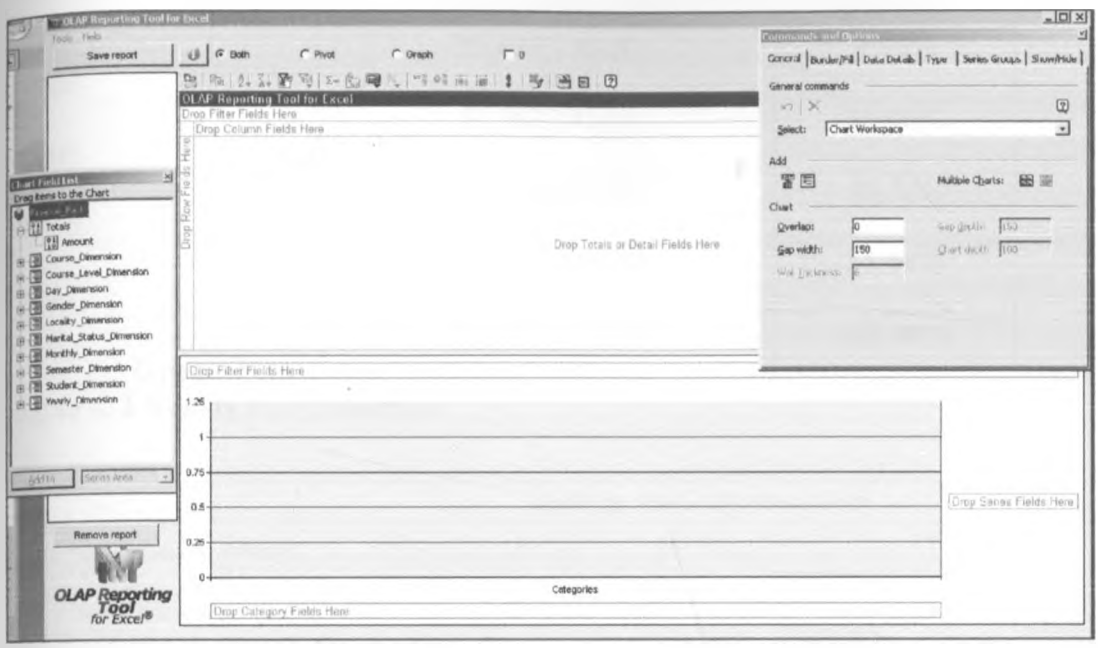


Figure 4. 12 OLAP Reporting Tool for Excel

Working environment consists of:

- A collection of integrated windows: OLAP Report Tool for Excel main window, Command and Option window and Chart Field List, designed to support a DSS
- Working Screen section - designed for dropping different fields depending report to be created
- OLAP Reporting Tool for Excel gets its data from OLAP database (data warehouse)

4.3 Analyzing and Reporting data in Data Warehouse database

4.3.1 Yearly Fee Collection

	Amount Collected	Percent
2008	27,475,780.00	9.75%
2009	59,382,245.00	21.06%
2010	95,034,075.00	33.71%
2011	100,043,259.00	35.48%
Grand Total	**281,935,359.00**	**100.00%**

Table 4. 1 Yearly Fee Collection

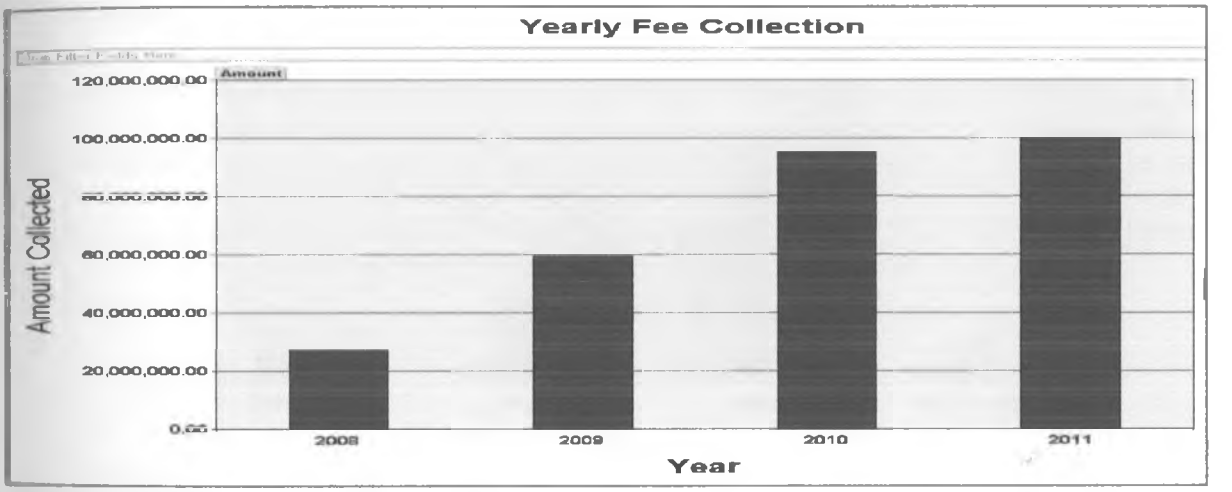


Figure 4. 13 Fees Collection per Year

There has been a steady increase in enrolment of students from 2009 following upgrading of the college from certificate or Diploma College to a constituent university college in 2008. This was evident in the amount revenue collected in the respective years. Revenue collection in 2009 and 2010 went up by 11.32% and 12.65% respectively compared to the previous respective years. However, there was sharp drop in 2011, constituting 1.78% although still more revenue was collected compared to the previous year

4.3.2 Fee Collection per Course per Year

	Certificate		Degree		Diploma		Higher Diploma	
	Amount	Percent	Amount	Percent	Amount	Percent	Amount	Percent
2008	748,567	2.95%	2,448,740	3.37%	21,853,923.	14.01%	2,424,550	8.72%
2009	7,593,106	29.88%	11,811,355	16.25%	33,046,219	21.18%	6,931,565	24.94%
2010	9,545,938	37.56%	33,530,481	46.13%	43,745,542	28.04%	8,212,114	29.54%
2011	7,526,186	29.61%	24,898,093	34.25%	57,391,060	36.78%	10,227,920	36.80%
Total	25,413,797	100.00%	72,688,669	100.00%	156,036,744	100.00%	27,796,149	100.00%

Table 4. 2 Fee Collection per Course per Year

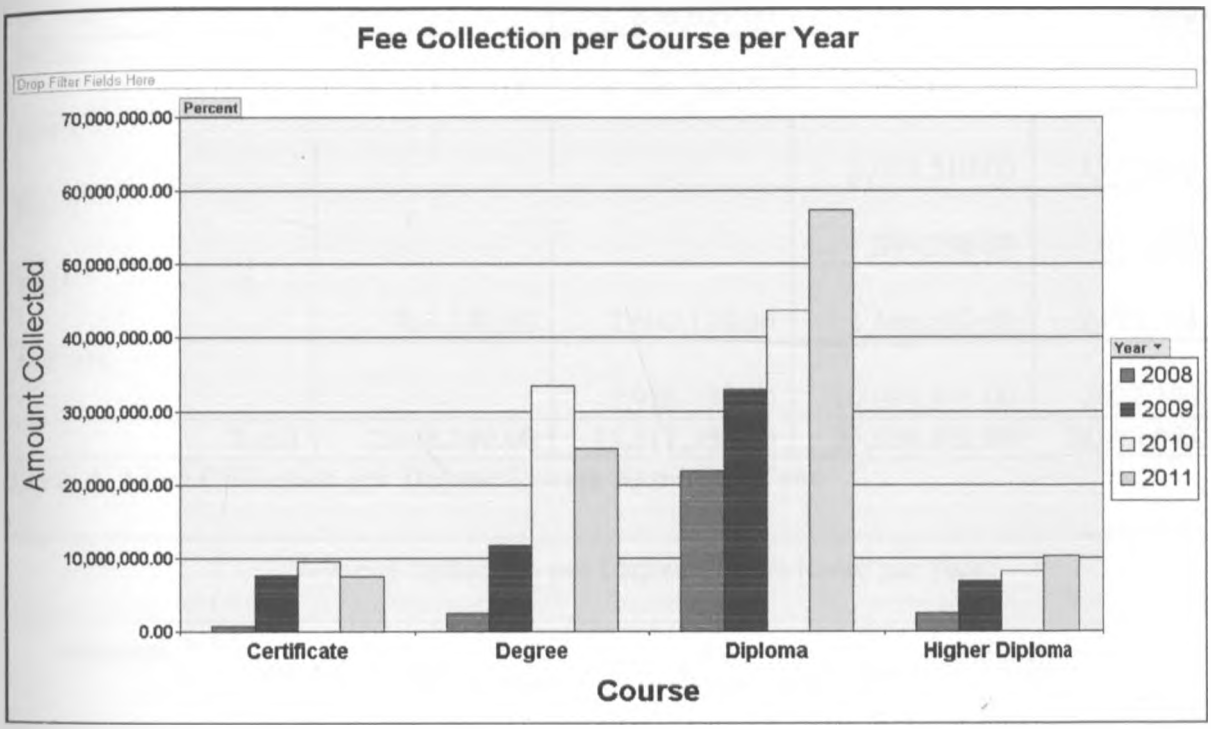


Figure 4. 14 Fee Collection per course per year

There has been constant increase in revenue collection from diploma and higher diploma courses throughout the four year period. It is important to note at this point that Diploma courses are the key income generators to the college before and even after the college was elevated. Both certificate and degree courses showed increase in revenue in 2009 and 2010 but a drop in 2011.

4.3.3 Fee Collection per Degree Course Name per Year

The college only offered one degree course - Bachelor of Computing and Information systems (BCIS) before it was elevation. This particular course was slowly replaced with Bachelor of Business Information Technology. Bachelor of Science (Information Technology) was introduced immediate after the college was upgraded. A year later, two more courses were introduced namely Bachelor of Commerce and Bachelor of Science

(Telecommunication and Information Engineering). The following year, two other courses were introduced that is Bachelor of Film Production and Animation, and Bachelor of Science (Computer Technology). All degree courses offered in 2009 registered an increase in revenue collection in 2010. However, in 2011, only Bachelor of Film and Animation course recorded an increase in revenue collection, while the rest registered a drop.

Course Name	2008	2009	2010	2011
BBIT		5,216,560.00	5,947,000.00	3,220,330.00
BCOM		836,627.00	10,061,056.00	7,772,659.00
BCIS	1,687,400.00	1,181,300.00	52,000.00	
BFPA			2,628,510.00	3,777,640.00
BSCT			881,398.00	605,800.00
BSIT	761,340.00	2,600,129.00	3,866,562.00	2,499,394.00
BSCTIE		1,976,739.00	10,093,955.00	7,022,270.00
Total	**2,448,740.00**	**11,811,355.00**	**33,530,481.00**	**24,898,093.00**

Table 4. 3 Fee Collection per Degree Course Name per Year

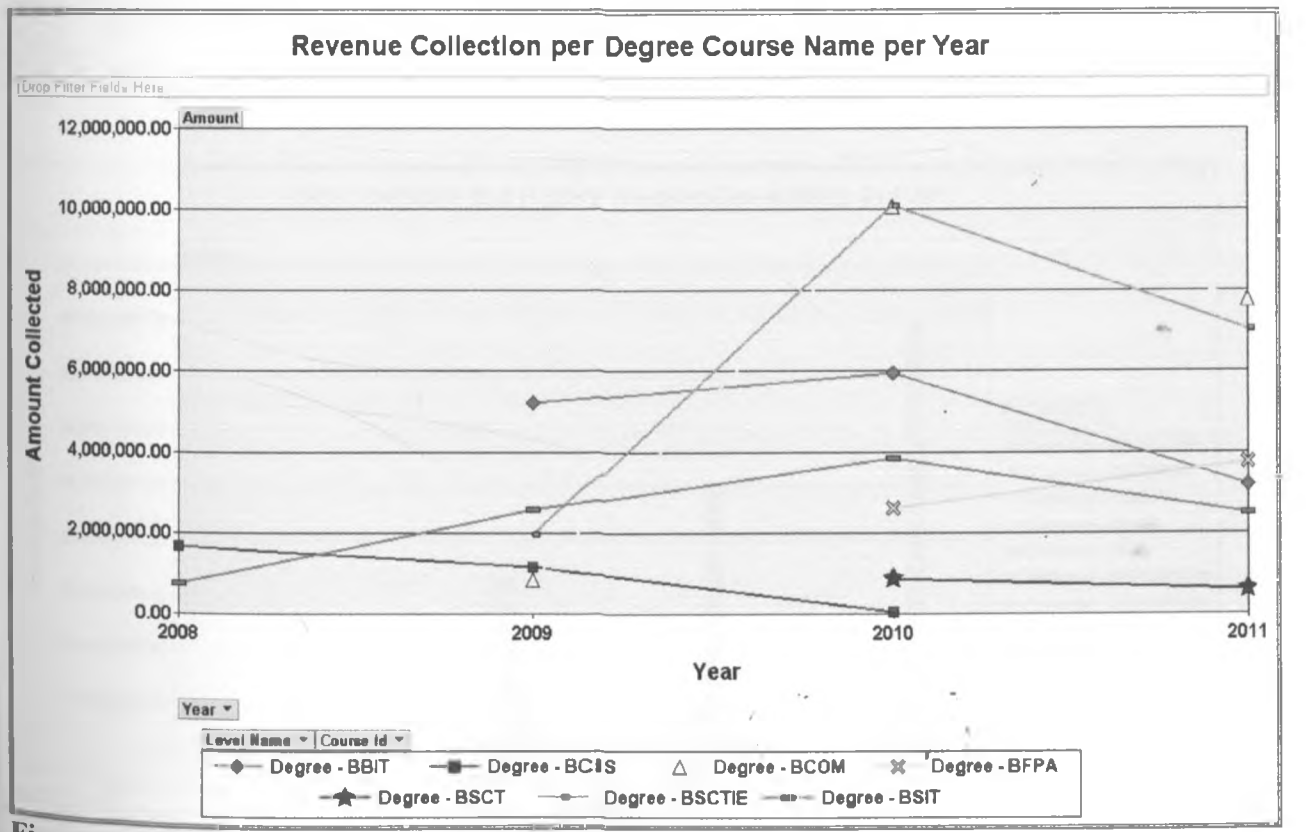


Figure 4. 15 Revenue Collection Per Degree Course Name per Year

4.3.4 Most Reliable and Highly Income Generating Course

Diploma in Mass Communication generates the highest revenue (33.74%) for the college each year for the last four years and there has been steady increase in revenue collection throughout the same period. Second in revenue generation is International Diploma in Computer Studies (13.01%), although it registered slight drop in 2010. The third is International Advanced Diploma in Computer Studies (9.54%). Diploma in Film Production and Animation is a course that has also record steady increase in revenue since its introduction in 2009. It is important to note that

4.3.4.1 Most Reliable and Highly Income Generating Course

Course	2008	2009	2010	2011	Total
AD PSUPPLY		415,400	8,800		424,200
CISA		74,000			74,000
CIPS4		2,085,752	2,872,857	916,860	5,875,469
CGADATS		214,810	724,520	379,920	1,319,250
CGDAT	27,900	809,520	375,390	172,700	1,385,510
DFPDN&ANI		1,088,330	4,586,170	5,558,580	11,233,080
DMASSCOM	13,064,300	16,721,923	25,239,830	40,090,488	95,116,541
IMIS2		119,100			119,100
IMISF		1,862,187	1,385,300	570,100	3,817,587
IDCS	8,761,723	9,655,197	8,552,675	9,702,412	36,672,007
Total	21,853,923	33,046,219	43,745,542	57,391,060	156,036,744

Table 4. 4 Most Reliable and Highly Income Generating Course

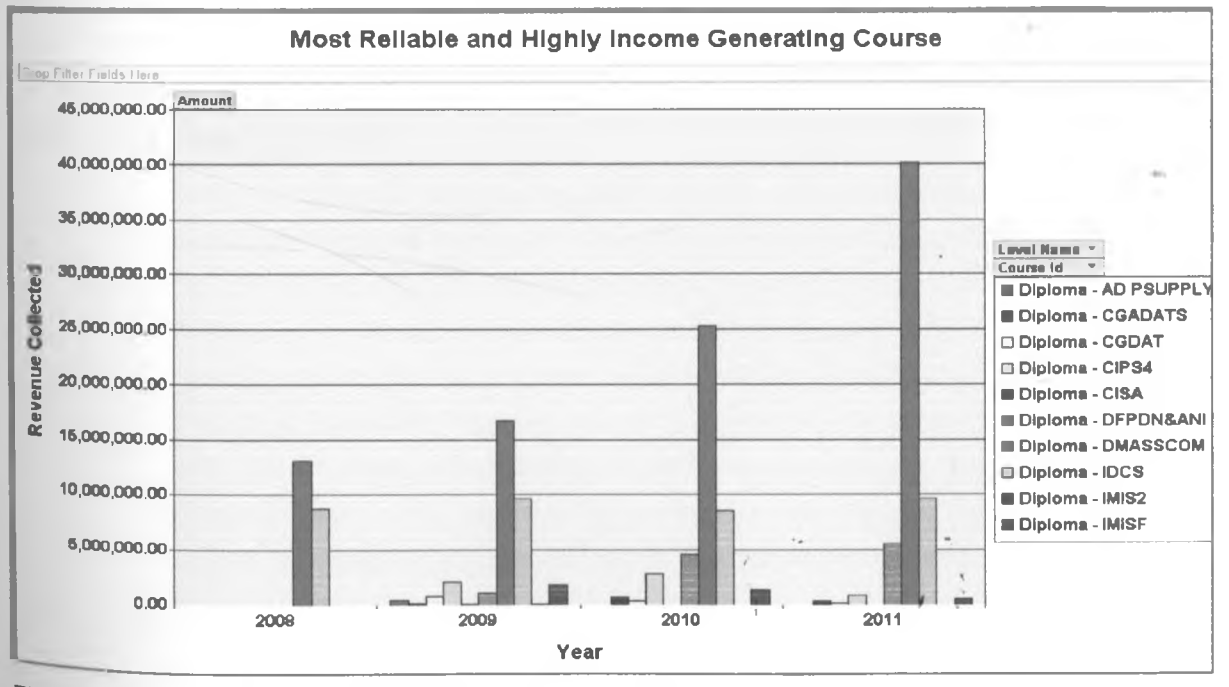


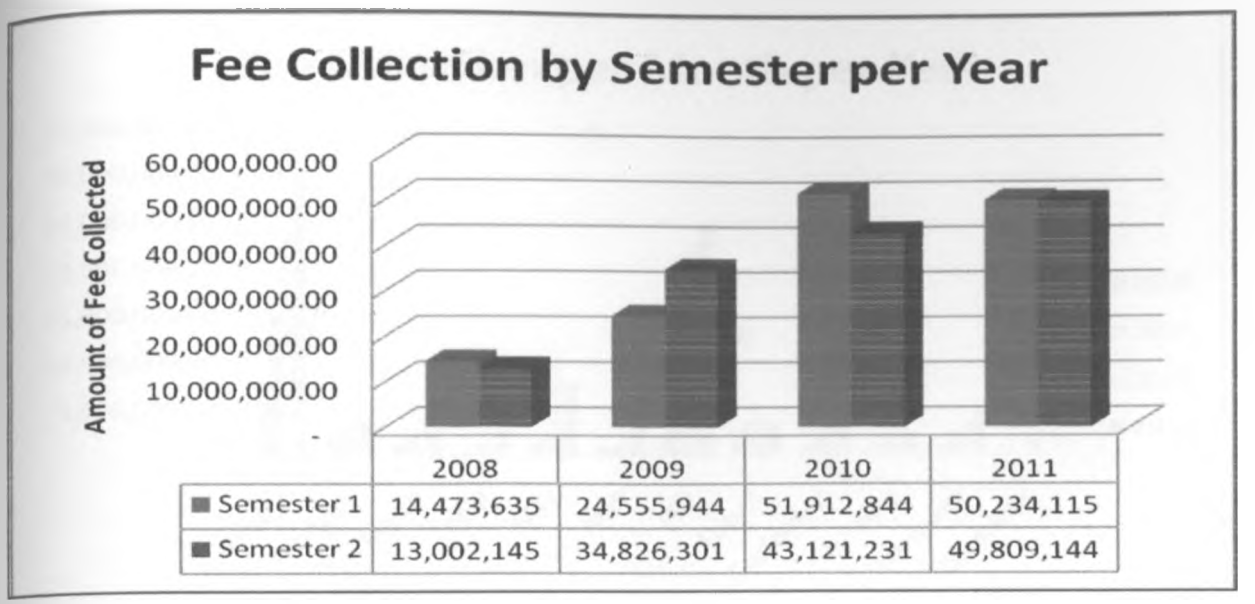
Figure 4. 16 Most Reliable and Highly Income generating Course

4.3.4.2 List of all courses offered at the college

Course Name	2008	2009	2010	2011	Grand Total	%
DMASSCOM	13,064,300	16,721,923	25,239,830	40,090,488	95,116,541	33.74
IDCS	8,761,723	9,655,197	8,552,675	9,702,412	36,672,007	13.01
IADCS	2,424,550	6,339,659	7,957,639	10,227,920	26,949,768	9.56
BSCTIE		1,976,739	10,093,955	7,022,270	19,092,964	6.77
BCOM		836,627	10,061,056	7,772,659	18,670,342	6.62
BBIT		5,216,560	5,947,000	3,220,330	14,383,890	5.10
DFPDN&ANI		1,088,330	4,586,170	5,558,580	11,233,080	3.98
BSIT	761,340	2,600,129	3,866,562	2,499,394	9,727,425	3.45
CMASSCOM	748,567	2,627,800	2,469,750	3,412,780	9,258,897	3.28
BFPA			2,628,510	3,777,640	6,406,150	2.27
CIPS4		2,085,752	2,872,857	916,860	5,875,469	2.08
CCNA1		1,497,067	1,809,171	1,000,800	4,307,038	1.53
BCIS	1,687,400	1,181,300	52,000		2,920,700	1.04
CIPS1		687,604	1,374,275	789,840	2,851,719	1.01
CIPS		741,535	1,046,910	534,646	2,323,091	0.82
CIPS L3		468,200	834,500	1,007,450	2,310,150	0.82
BSCT			881,398	605,800	1,487,198	0.53
CGDAT	27,900	809,520	375,390	172,700	1,385,510	0.49
CGADATS		214,810	724,520	379,920	1,319,250	0.47
CCNA2		450,100	648,100	158,060	1,256,260	0.45
CIC		266,600	375,102	381,710	1,023,412	0.36
CCNA3		303,900	483,500	152,500	939,900	0.33
CGCATS		342,300	232,130	35,900	610,330	0.22
CCNA4		208,000	272,500	52,500	533,000	0.19
CIPSGD		253,746	254,475		508,221	0.28
AD PSUPPLY		415,400	8,800		424,200	0.15
IMISDII		338,160			338,160	0.12
IMIS2		119,100			119,100	0.04
CISA		74,000			74,000	0.03
Grand Total	**27,475,780**	**59,382,245**	**95,034,075**	**100,043,259**	**281,935,359**	

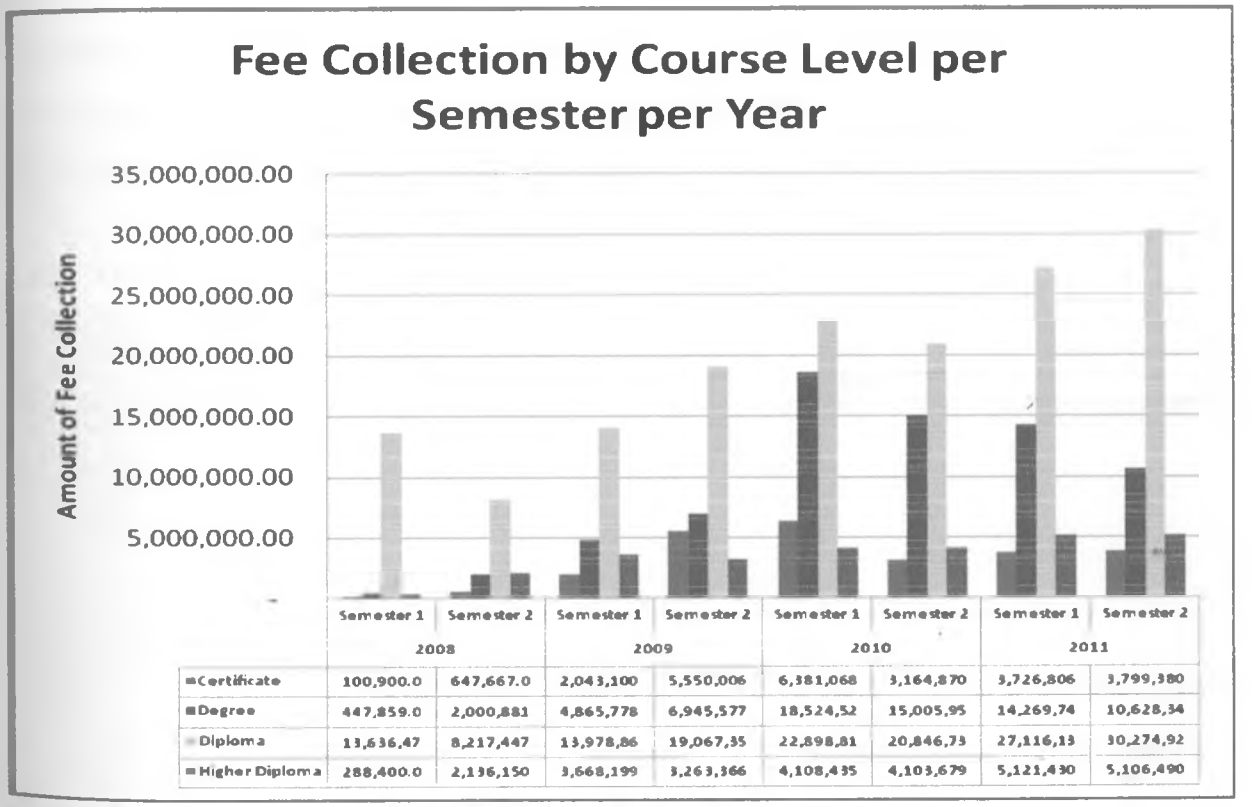
Table 4. 5 Course Hierarchy

4.3.5 Fee Collection by Semester per Year



	2008	2009	2010	2011
Semester 1	14,473,635	24,555,944	51,912,844	50,234,115
Semester 2	13,002,145	34,826,301	43,121,231	49,809,144

Figure 4. 17 Fee Collection per Semester per Year



	2008		2009		2010		2011	
	Semester 1	Semester 2	Semester 1	Semester 2	Semester 1	Semester 2	Semester 1	Semester 2
Certificate	100,900.0	647,667.0	2,043,100	5,550,006	6,381,068	3,164,870	3,726,806	3,799,380
Degree	447,859.0	2,000,881	4,865,778	6,945,577	18,524,52	15,005,95	14,269,74	10,628,34
Diploma	13,636,47	8,217,447	13,978,86	19,067,35	22,898,81	20,846,73	27,116,13	30,274,92
Higher Diploma	288,400.0	2,136,150	3,668,199	3,263,366	4,108,435	4,103,679	5,121,430	5,106,490

Figure 4. 18 Fee per Level per Semester per Year

4.3.6 Fee Collection per Month Year

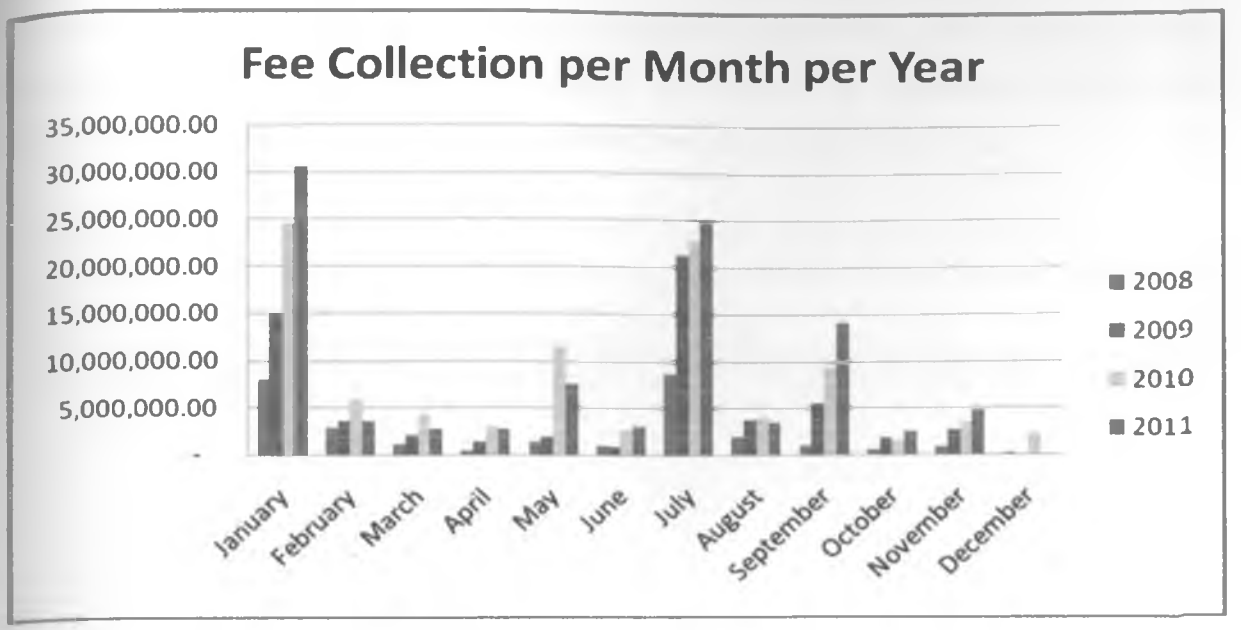


Figure 4. 19 Fee Collection per Month per Year

Drill Through

Drill Through enables you to retrieve, from the source system, the transactions from which the summary data in Analysis Services was created.

4.3.7 OLAP database enables you to slice and dice summarized data

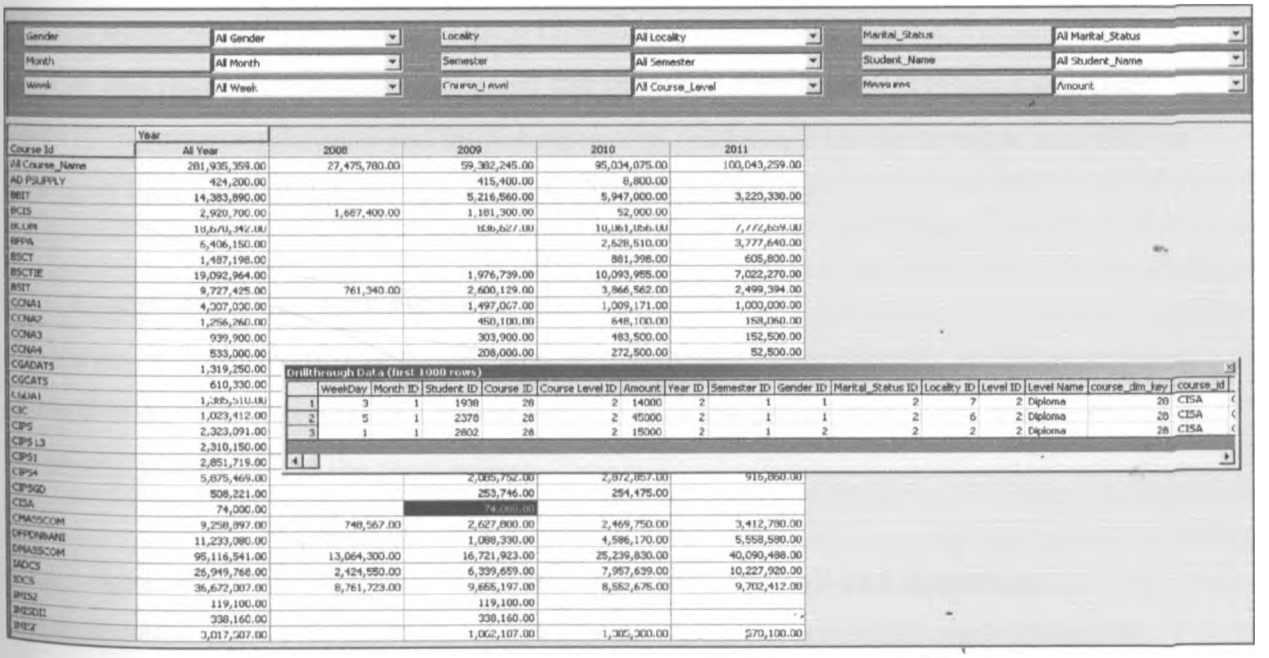
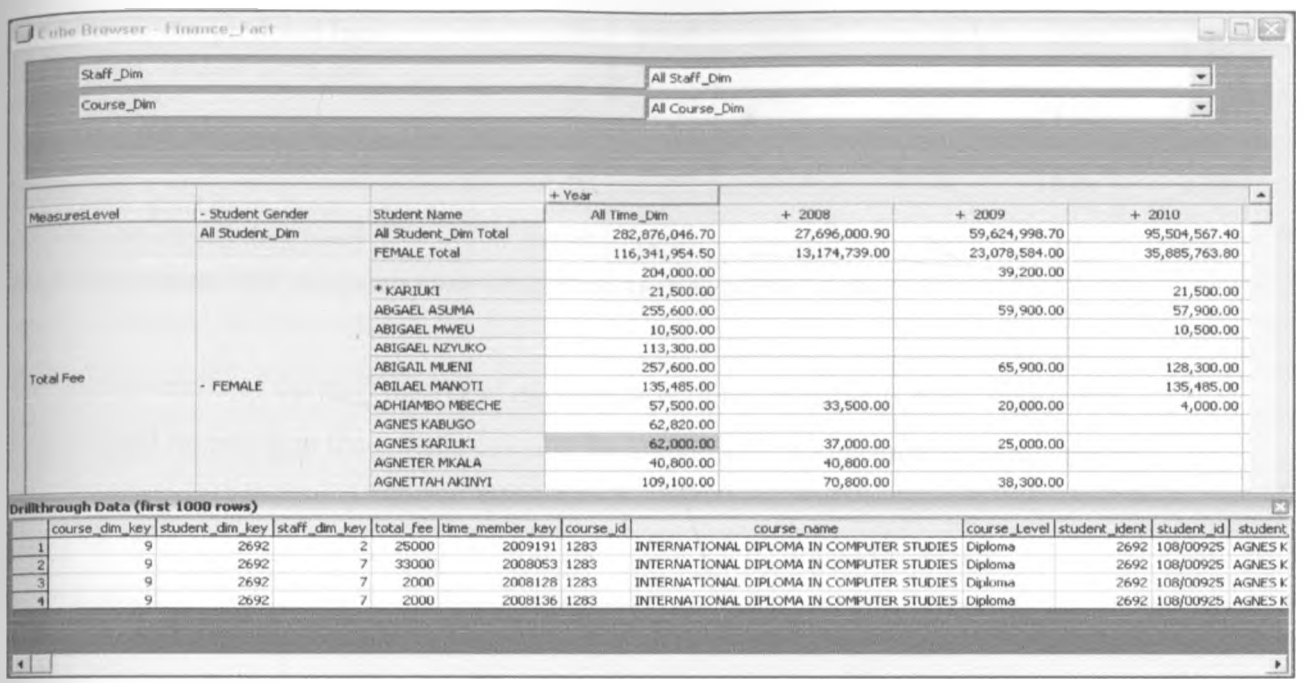


Figure 4. 20 Drill Down

The OLAP database enables you to slice and dice through summarized data. Data contained in such systems is summarized from the transactions contained in the relational source database. Summarized data in OLAP systems are helpful for conducting analysis and determining business information, such as trends, exceptions, errors and so on.


Cube Browser - Finance_Fact

Staff_Dim: All Staff_Dim

Course_Dim: All Course_Dim

MeasuresLevel	- Student Gender	Student Name	+ Year: All Time_Dim	+ 2008	+ 2009	+ 2010
	All Student_Dim	All Student_Dim Total	282,876,046.70	27,696,000.90	59,624,998.70	95,504,567.40
		FEMALE Total	116,341,954.50	13,174,739.00	23,078,584.00	35,885,763.80
			204,000.00		39,200.00	
		* KARIUKI	21,500.00			21,500.00
		ABGAEL ASUMA	255,600.00		59,900.00	57,900.00
		ABIGAEL MWEU	10,500.00			10,500.00
		ABIGAEL NZYUKO	113,300.00			
Total Fee		ABIGAIL MUENI	257,600.00		65,900.00	128,300.00
	- FEMALE	ABILAEL MANOTI	135,485.00			135,485.00
		ADHIAMBO MBECHE	57,500.00	33,500.00	20,000.00	4,000.00
		AGNES KABUGO	62,820.00			
		AGNES KARIUKI	62,000.00	37,000.00	25,000.00	
		AGNETER MKALA	40,800.00	40,800.00		
		AGNETTAH AKINYI	109,100.00	70,800.00	38,300.00	

Drillthrough Data (first 1000 rows)

	course_dim_key	student_dim_key	staff_dim_key	total_fee	time_member_key	course_id	course_name	course_Level	student_ident	student_id	student
1	9	2692	2	25000	2009191	1283	INTERNATIONAL DIPLOMA IN COMPUTER STUDIES	Diploma	2692	108/00925	AGNES K
2	9	2692	7	33000	2008053	1283	INTERNATIONAL DIPLOMA IN COMPUTER STUDIES	Diploma	2692	108/00925	AGNES K
3	9	2692	7	2000	2008128	1283	INTERNATIONAL DIPLOMA IN COMPUTER STUDIES	Diploma	2692	108/00925	AGNES K
4	9	2692	7	2000	2008136	1283	INTERNATIONAL DIPLOMA IN COMPUTER STUDIES	Diploma	2692	108/00925	AGNES K

Figure 4. 21 Drill Through

When a trend, an exception, or an error is identified, you can drill down to the transaction level to find the source data that generated this behavior. This feature provides a complete integration between relational and multidimensional database. You can navigate and analyze data from both types of databases in a single client environment.

4.4 Quantitative Research Findings

The survey interviewed different categories of stakeholders to gauge their understanding and perception on several aspects of customer service. Several departments were represented in the survey as indicated in the chart below.

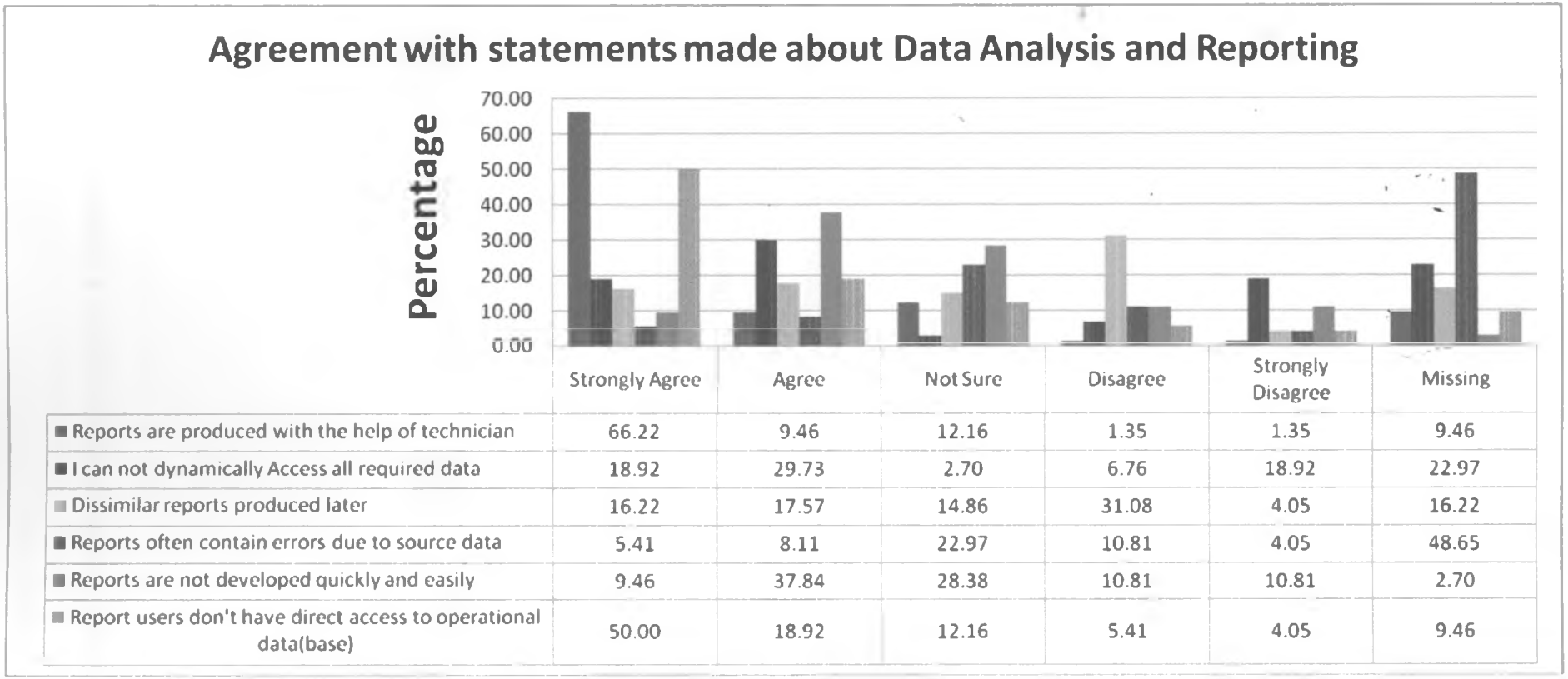
4.4.1 Agreement with statements made about Data Analysis and Reporting

The table below shows the staff's level of agreement with the statements made about data analysis and reporting in the college

	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree	Missing
Reports are produced with the help of technician	66.22	9.46	12.16	1.35	1.35	9.46
I can not dynamically Access all required data	18.92	29.73	2.70	6.76	18.92	22.97
Dissimilar reports produced later	16.22	17.57	14.86	31.08	4.05	16.22
Reports often contain errors due to source data	5.41	8.11	22.97	10.81	4.05	48.65
Reports are not developed quickly and easily	9.46	37.84	28.38	10.81	10.81	2.70
Report users don't have direct access to operational data(base)	50.00	18.92	12.16	5.41	4.05	9.46

Table 4. 6 Agreement with statements made about Data Analysis and Reporting

The above results of the staff's level of agreement with the statements made about data analysis and reporting in the college can also be graphically represented as shown below:



	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree	Missing
Reports are produced with the help of technician	66.22	9.46	12.16	1.35	1.35	9.46
I can not dynamically Access all required data	18.92	29.73	2.70	6.76	18.92	22.97
Dissimilar reports produced later	16.22	17.57	14.86	31.08	4.05	16.22
Reports often contain errors due to source data	5.41	8.11	22.97	10.81	4.05	48.65
Reports are not developed quickly and easily	9.46	37.84	28.38	10.81	10.81	2.70
Report users don't have direct access to operational data(base)	50.00	18.92	12.16	5.41	4.05	9.46

Figure 4. 22 Agreement with the statements made about data analysis and reporting

When asked their level of agreement with some statements regarding the generation of reports and data analysis, majority of the staff (66.22%) strongly agree that report users depend on IT professionals in producing report. Only 2.70% argue that they can create a report on their own. Also, half of the decision makers in the college (50.00%) strongly agree that they understand what data means, but they don't have easy access to the data. Furthermore, the respondents are in agreement (47.30%) that reports are not developed quickly and easily although a smaller percentage (13.52%) acknowledges that any typical report is filled with errors due to the source data.

4.4.2 How long it takes to create a report

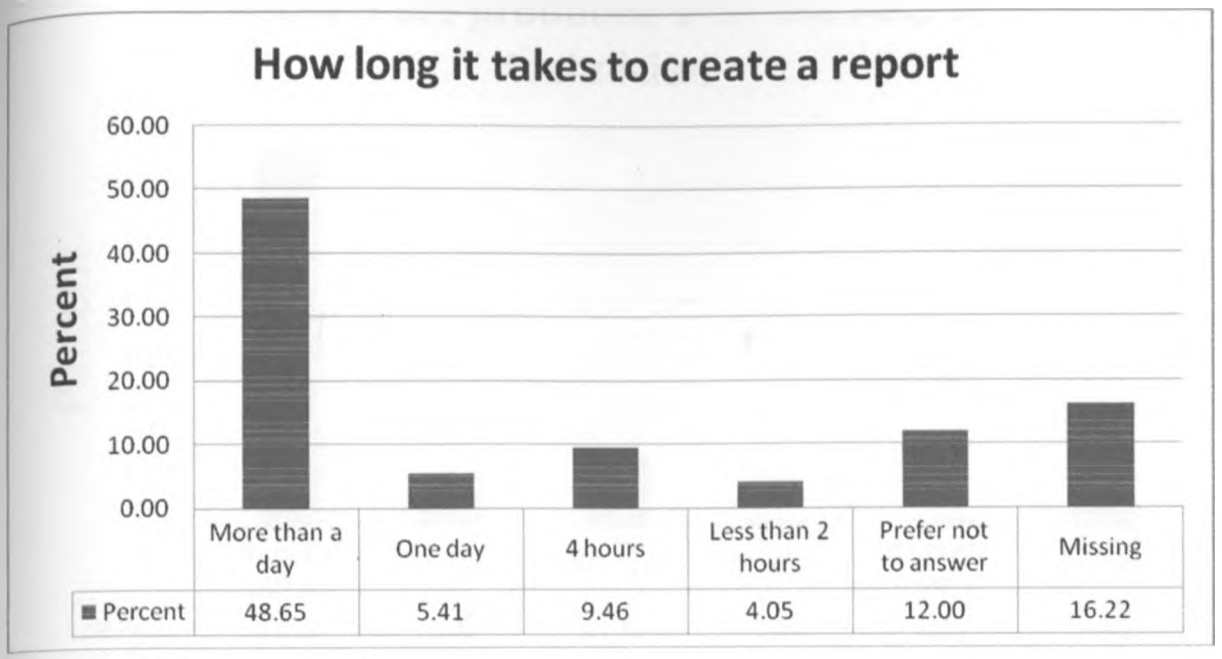


Figure 4. 23 How long it takes to create a report

Most of the staff (54.06%) acknowledges that the typical report takes at least a day to create as opposed to 13.6% who say that it takes at most 4 hours. There is however a significant number of staff (32.4%) who preferred not to respond or ignored the question. This study has revealed that users need a user friendly interface to easily turn the analysis into a report.

4.4.3 Reports are produced with the help of technician

	Frequency	Percent
Strongly Agree	49	66.2
Agree	7	9.5
Not Sure	9	12.2
Disagree	1	1.4
Strongly Disagree	1	1.4
Missing	7	9.5
Total	**74**	**100.0**

Table 4. 7 Reports are produced with the help of Technician

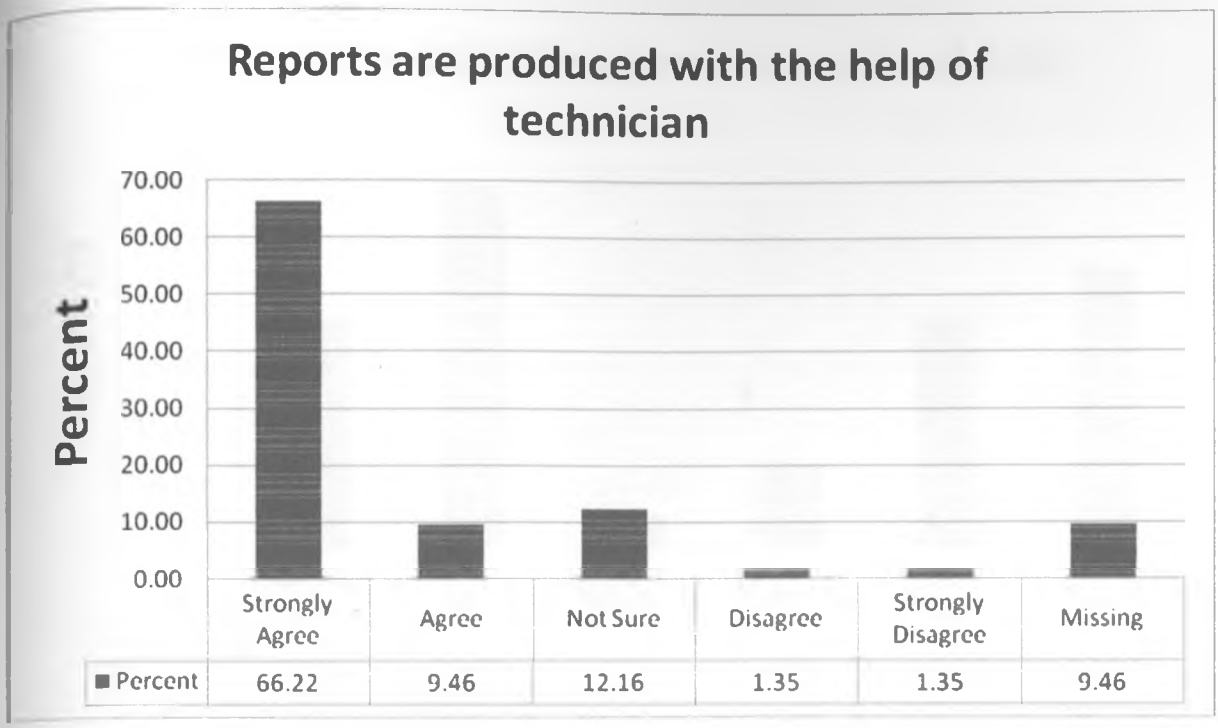


Figure 4. 24 Reports are produced with the help of Technician

75.68% of the respondents are in agreement that reports are produced at least with the help of a technician. Only 2.7% argue that they can create a report on their own while 12.16% are not sure. 9.46% felt there was no need to respond. It is important to note that there is an urgent need to break the dependency and make departments independence from IT department.

4.4.4 I cannot dynamically Access all required data

	Frequency	Percent
Strongly Agree	14	18.9
Agree	22	29.7
Not Sure	2	2.7
Disagree	5	6.8
Strongly Disagree	14	18.9
Missing	17	23.0
Total	**74**	**100.0**

Table 4. 8 I cannot dynamically Access all required data

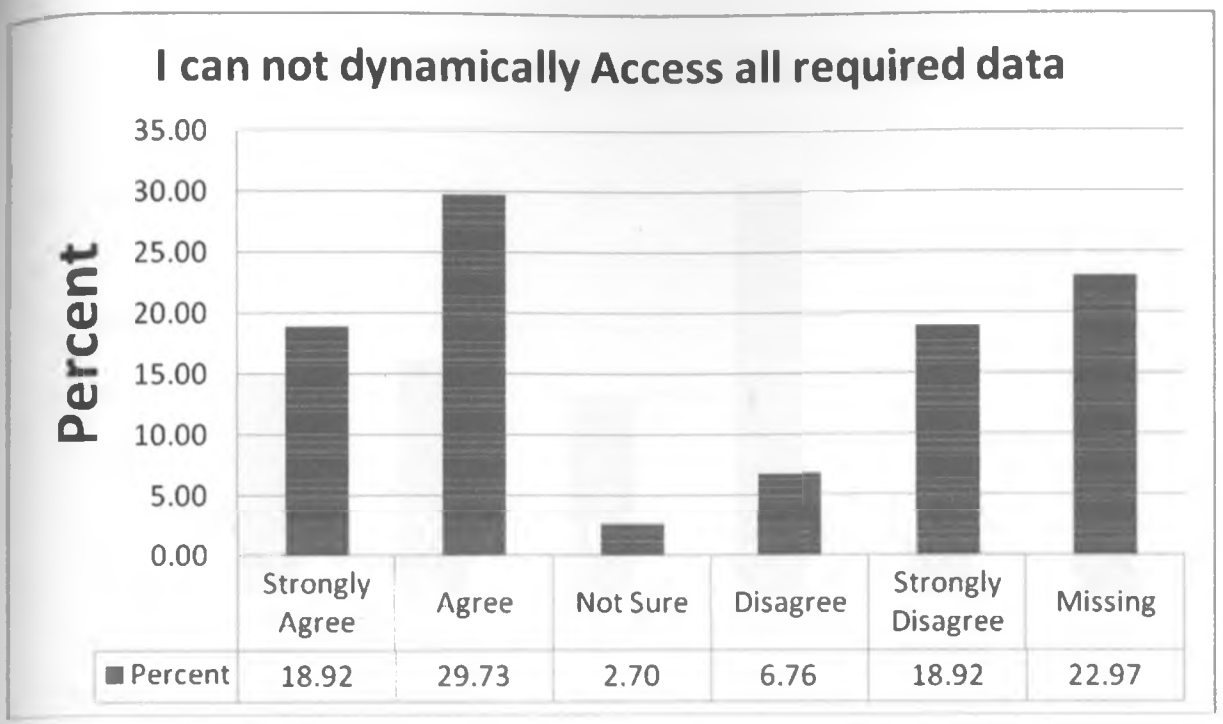


Figure 4. 25 I cannot dynamically Access all required data

The survey has shown that 48.6% strongly argue that they need dynamic access to all the data for them do any serious analysis while 25.7% say they are satisfied the way they interact with data. More importantly, users need an easy way to transform their analysis into report.

4.4.5 Dissimilar reports produced later

	Frequency	Percent
Strongly Agree	12	16.2
Agree	13	17.6
Not Sure	11	14.9
Disagree	23	31.1
Strongly Disagree	3	4.1
Missing	12	16.2
Total	**74**	**100.0**

Table 4. 9 Dissimilar reports produced later

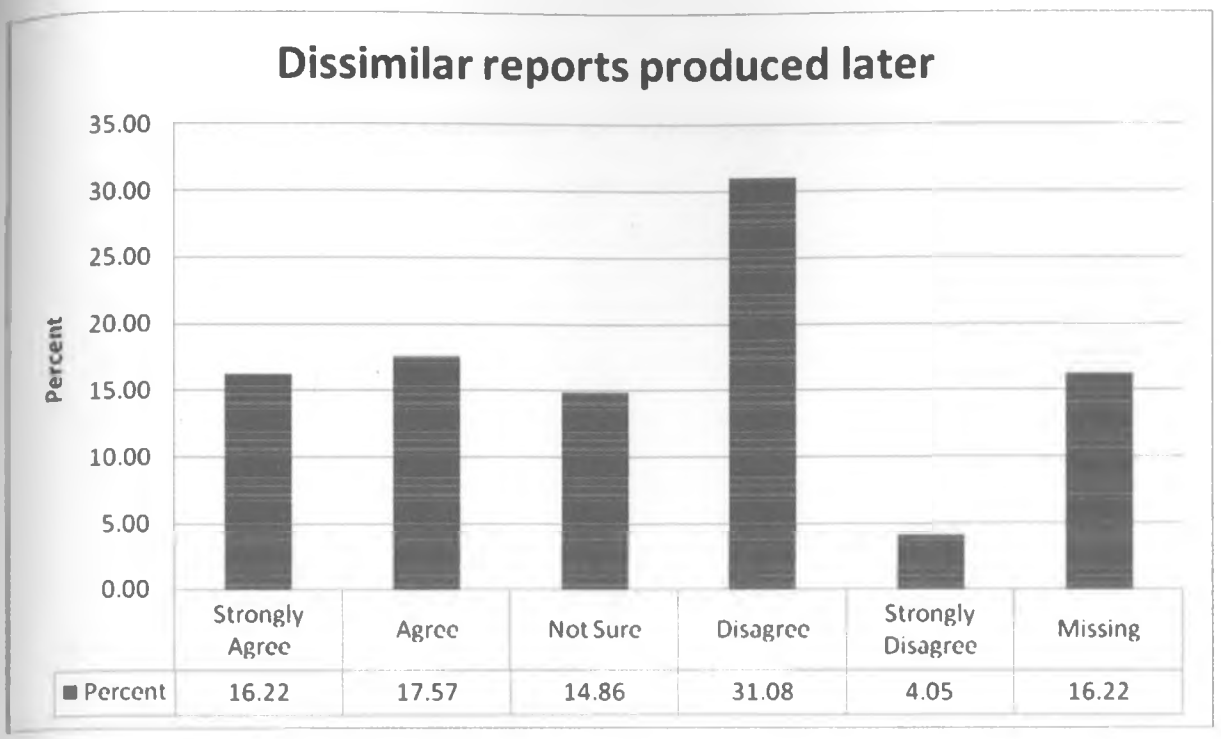


Figure 4. 26 Dissimilar reports produced later

Many users are accustomed to working with many versions of truth. 33.79% of the staff disagree and claim that similar reports are generated several months later against similar data. However, 35.13% claim that they have been creating dissimilar reports due to frequent change of data and therefore sometimes they are forced to maintain a spreadsheet database with each month's data. According to this group, there is a need for one version of the truth so that same reports are generated at different times by different people

4.4.6 Reports often contain errors due to source data

	Frequency	Percent
Strongly Agree	4	5.4
Agree	6	8.1
Not Sure	17	23.0
Disagree	8	10.8
Strongly Disagree	3	4.1
Missing	36	48.6
Total	**74**	**100.0**

Table 4. 10 Reports often contain errors due to source data

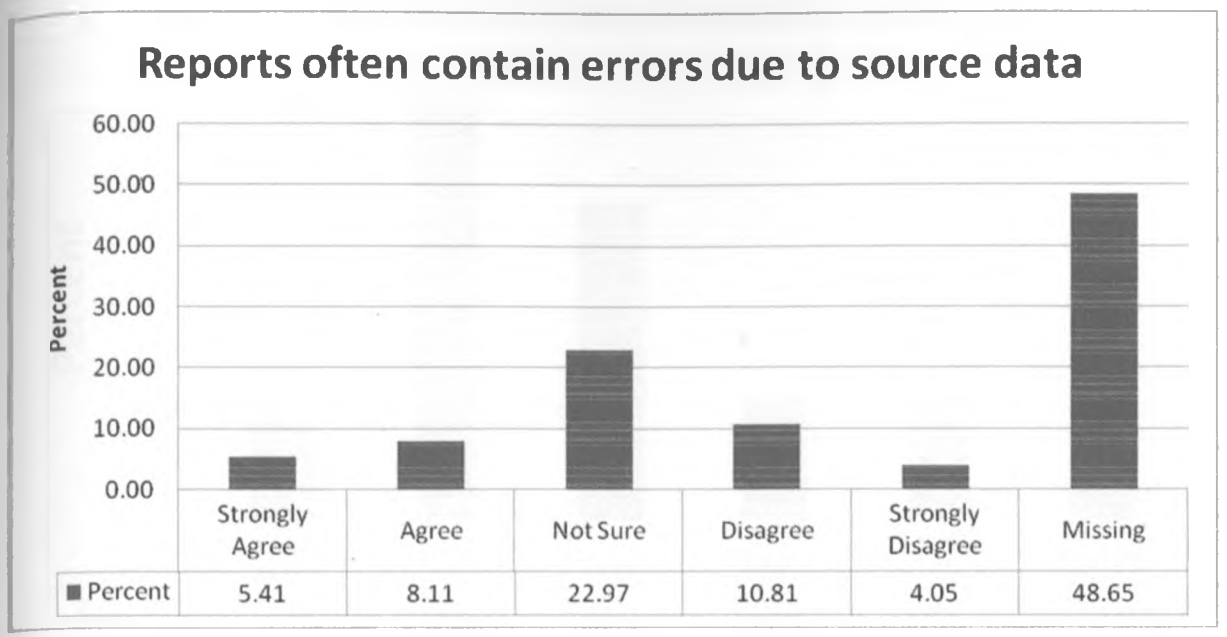


Figure 4. 27 Reports often contain errors due to source data

13.52% agree that reports often use data from spreadsheet database, many of which contain errors although 14.86% claim that reports created are with much less possibility of errors. Typical report generators hardly contain practical way to reconcile the results to an ultimate truth.

4.4.7 Reports are NOT developed quickly and easily

	Frequency	Percent
Strongly Agree	7	9.5
Agree	28	37.8
Not Sure	21	28.4
Disagree	8	10.8
Strongly Disagree	8	10.8
Missing	2	2.7
Total	**74**	**100.0**

Table 4. 11 Reports are not developed quickly and easily

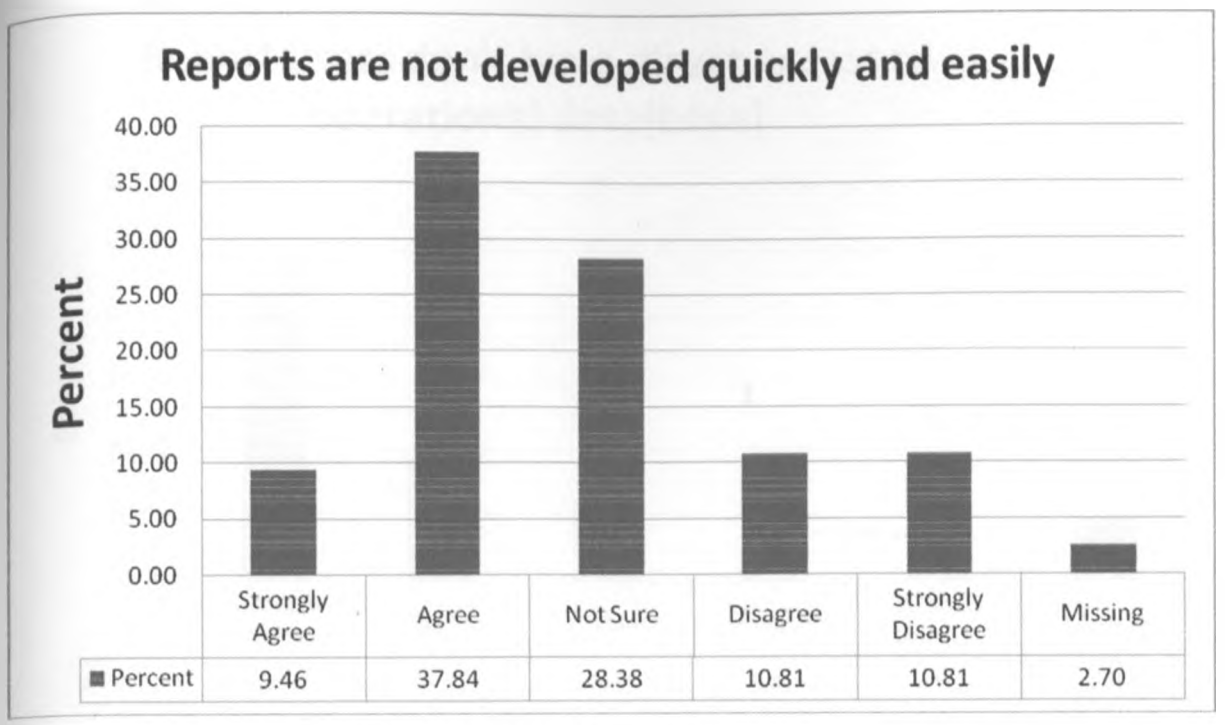


Figure 4. 28 Reports are NOT Developed Quickly and Easily

47.3% articulate that they have to work long hours to create a report while 21.6% say that new reports can be created in a few minutes or hours

4.4.8 Report users don't have direct access to operational data(base)

	Frequency	Percent
Strongly Agree	37	50.0
Agree	14	18.9
Not Sure	9	12.2
Disagree	4	5.4
Strongly Disagree	3	4.1
Missing	7	9.5
Total	**74**	**100.0**

Table 4. 12 Report users don't have direct access to operational data(base)

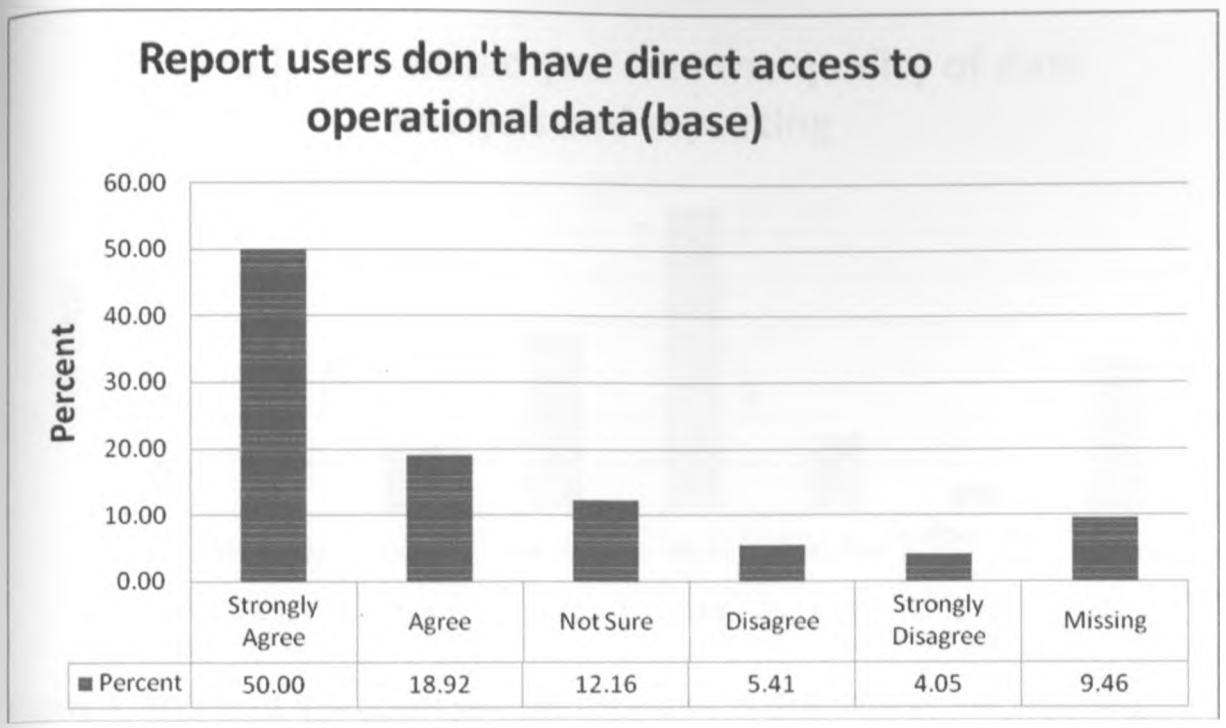


Figure 4. 29 Report users don't have direct access to operational data(base)

Most of the staff (68.92%) was dissatisfied with the level of interaction with data. Only 9.46% show satisfaction

4.4.9 Overall, how would you rate the quality of data analysis and reporting?

	Frequency	Percent
Very Good	11	14.9
Good	5	6.8
Not Sure	14	18.9
Poor	24	32.4
Very Poor	6	8.1
Prefer not to answer	2	2.7
Missing	12	16.2
Total	**74**	**100.0**

Table 4. 13 Overall, how would you rate the quality of data analysis and reporting

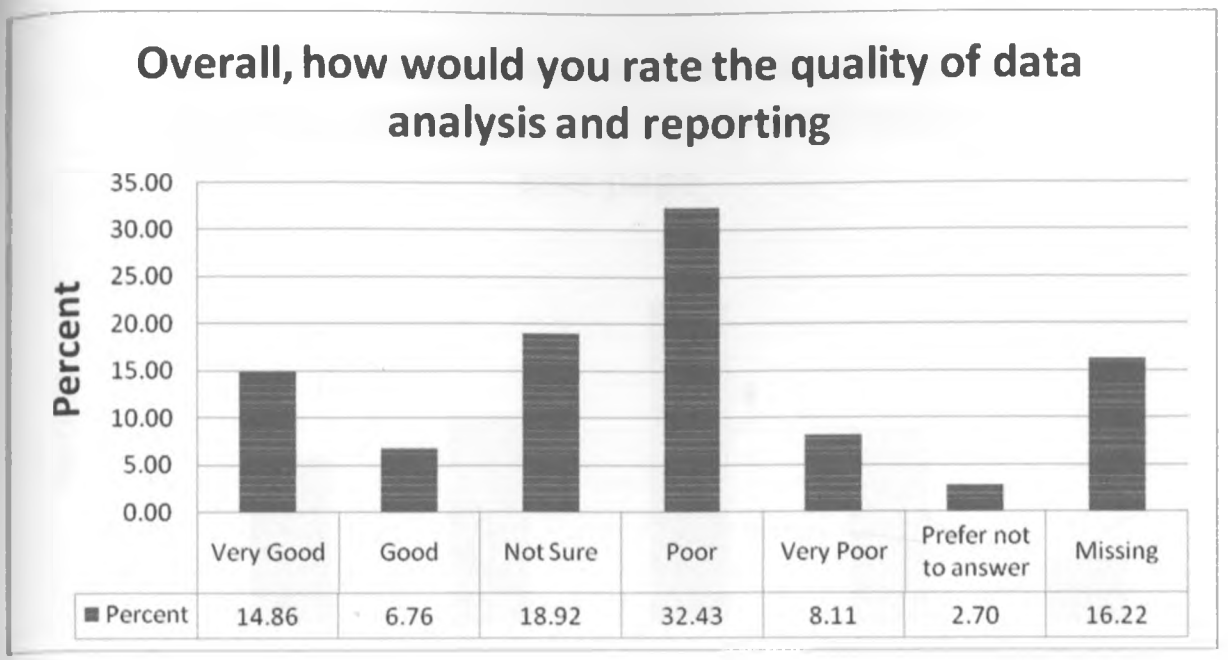


Figure 4. 30 Overall, how would you rate the quality of data analysis and reporting?

Only 21.7% of the respondents feel satisfied with the quality of data analysis and reporting. 40.5% rate it poorly and feel very dissatisfied but 37.8% fall in between.

4.4.10 What is your overall satisfaction level of analyzing and reporting data from any number of sources on one page?

	Frequency	Percent
Satisfied	13	17.6
Not Sure	16	21.6
Dissatisfied	25	33.8
Prefer Not to Answer	15	20.3
Missing	5	6.8
Total	**74**	**100.0**

Table 4. 14 Overall satisfaction level of analyzing and reporting data

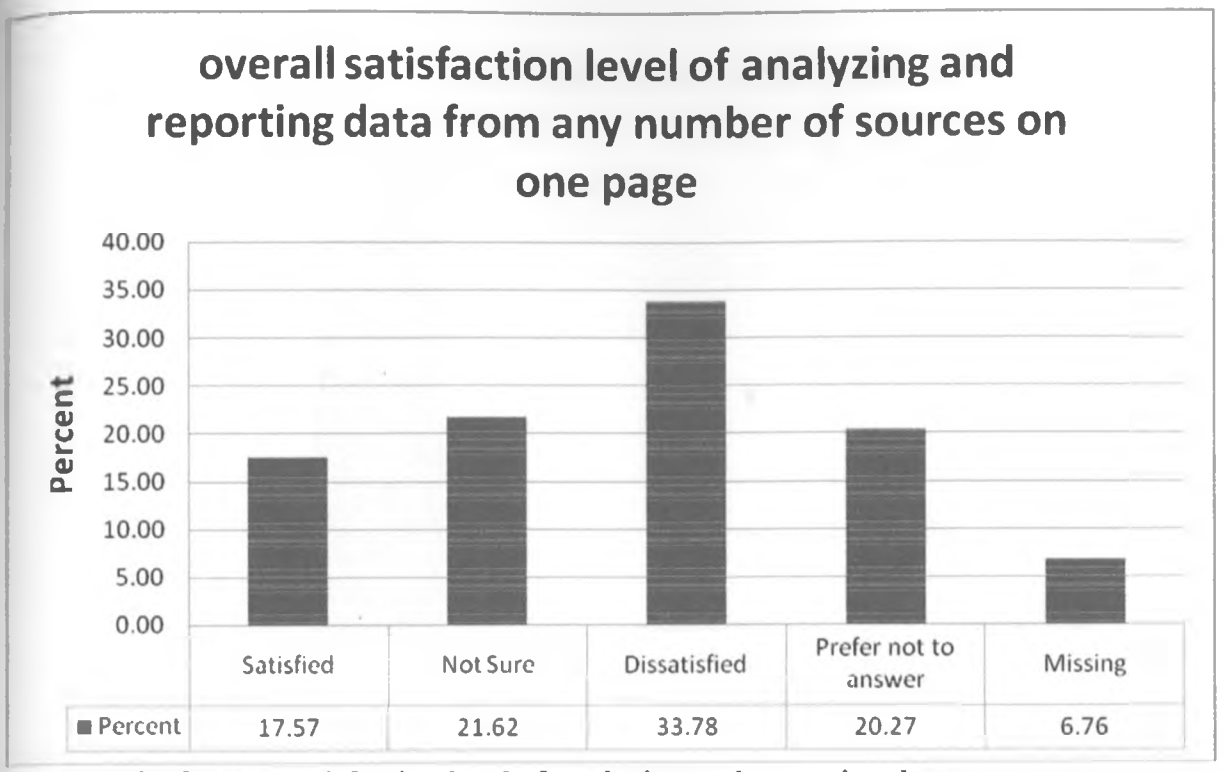


Figure 4. 31 Overall satisfaction level of analyzing and reporting data

On overall satisfaction level of analyzing and reporting data from any number of source on one page, only 17.5% were satisfied though 33.8 felt dissatisfied but still a majority of the respondents (48.7%) were not sure or ignored question

4.4.11 Have you used Ms Excel before

	Frequency	Percent
Yes	69	93.2
No	1	1.4
Missing	4	5.4
Total	**74**	**100.0**

Table 4. 15 Have you used Ms Excel

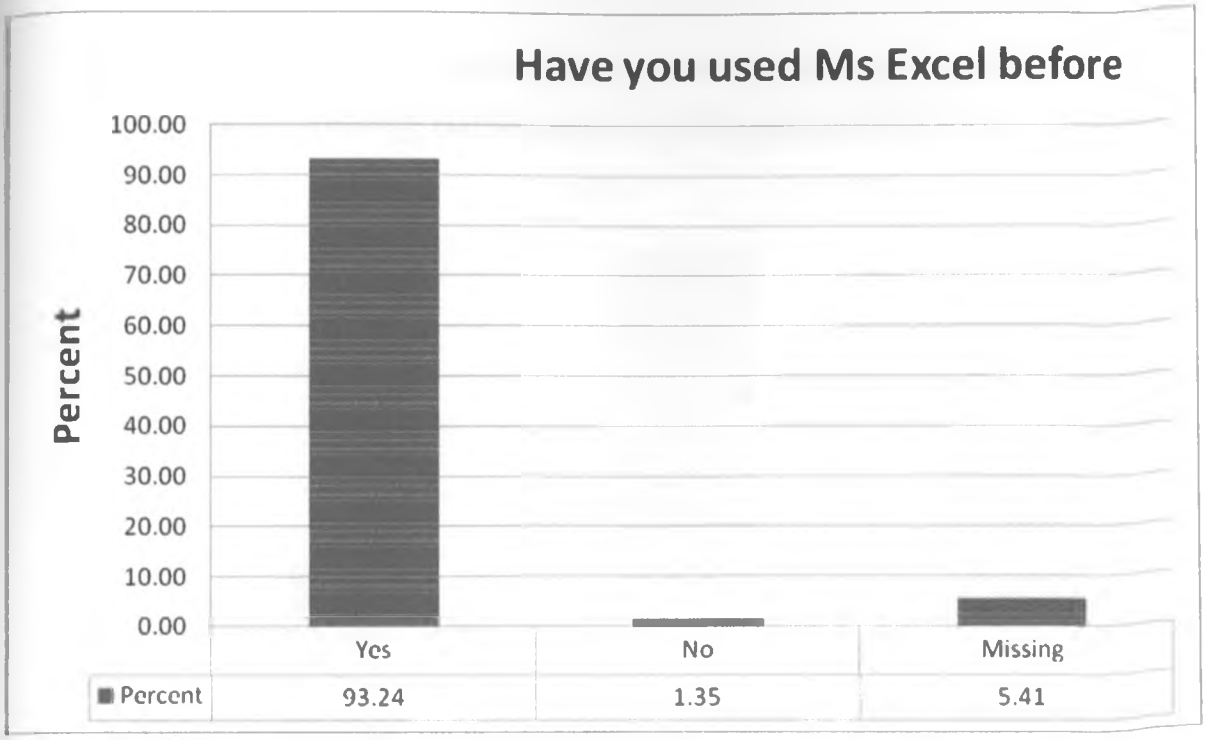


Figure 4. 32 Have you used Ms Excel before

Most of the staff is familiar Ms Excel. 93.2% say that they have used Ms Excel before with only 1.4% saying that they are not.

4.4.12 In the past, have you used Ms Excel as an outstanding Business Intelligence (BI) tool?

	Frequency	Percent
Yes	6	8.1
No	44	59.5
Missing	24	32.4
Total	**74**	**100.0**

Table 4. 16 Have you used Excel as BI

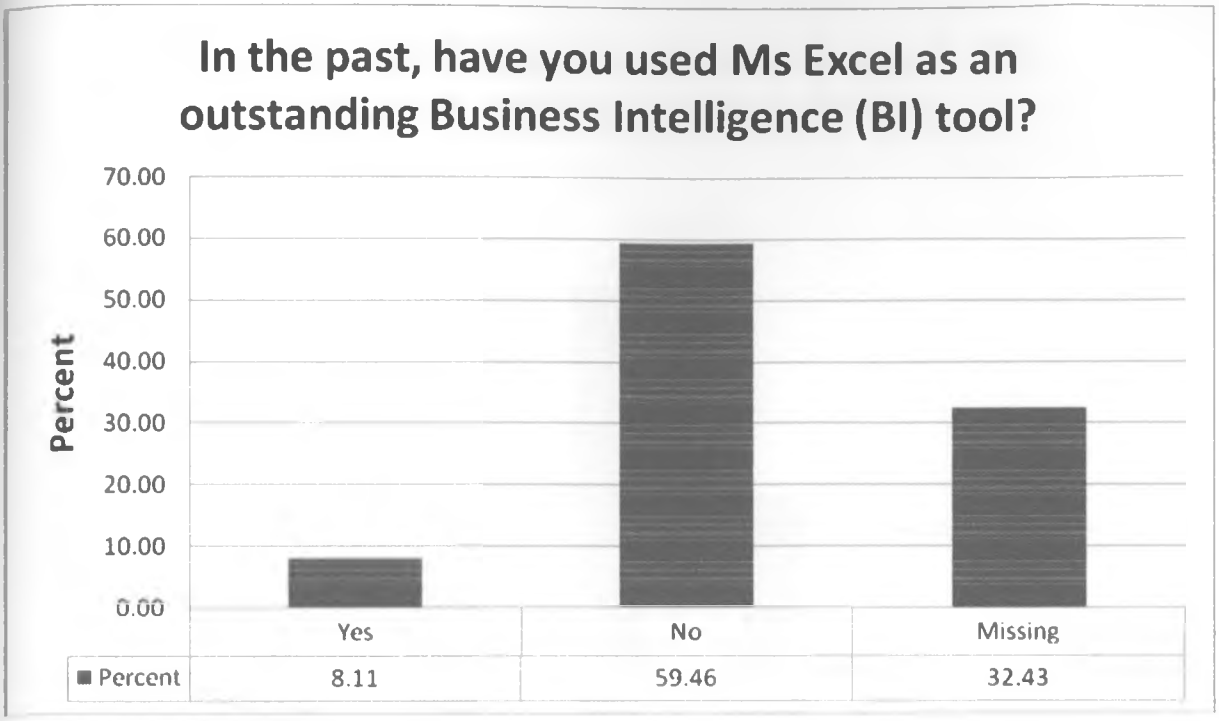


Figure 4. 33 Have you used Excel as BI

More than half of the staff (59.46%) have not used Ms Excel as an outstanding Business Intelligence (BI) tool? while 8.11% believe they have done so. The rest ignore to answer the question

4.4.13 Ms Excel can become a key component of a powerful and inexpensive BI tool

	Frequency	Percent
Definitely	9	12.2
Probably	13	17.6
Not Sure	30	40.5
Probably Not	17	23.0
Prefer not to Answer	3	4.1
Missing	2	2.7
Total	**74**	**100.0**

Table 4. 17 Ms Excel can be BI tool

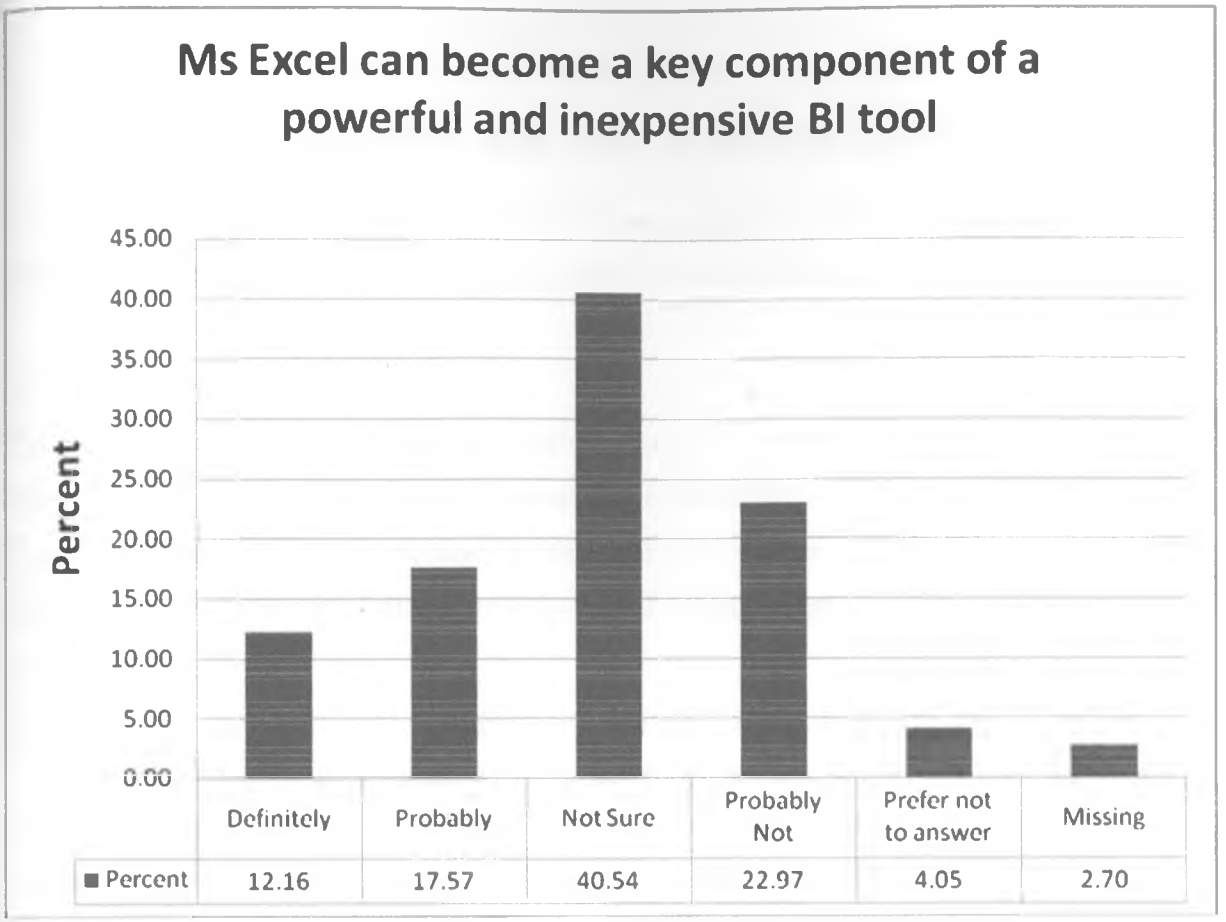


Figure 4. 34 Ms Excel can be BI tool

There is a general feeling (29.73%) that Ms Excel can become an outstanding Business Intelligence system. 22.97% claim that it won't be possible but still 47.29% have no idea about the same.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This chapter provides an overview of the report, describes how research objectives were addressed, a summary of research findings and recommendations.

5.1 Overview

This report began with a broad aim which sought to establish the level of dependency between top decision makers and IT personnel in accessing, analyzing and reporting data. The project also sought to determine the most valuable way to analyze data quickly and turn analysis into a report with little or no help from IT department. Specific research objectives were:

1 To determine the level of dependency between decision makers and IT professionals in the college in accessing, analyzing and reporting data

2 To establish the most valuable technology for the college users to dynamically access, analyze and report data easily and quickly

3 To design and develop a single multidimensional-cube (data warehouse) database

4 To implement an On Line Analytical Processing (OLAP) program that provides all the key people in the college with access to whatever level of information needed for the college to survive and prosper in an increasingly competitive world.

5.2 How Research Objectives were Achieved

This study outlined four research questions. Each question addressed respective research objective in the light of the analysis undertaken in this report. Research questions were as follows:

1. What managerial gaps are there which hinder decision making process in the constituent college?
2. What is the most valuable technology for the college users in accessing, analyzing and reporting data easily and quickly?
3. How will the technology be implemented to narrow the managerial gaps in the college?
4. How can the technology be optimized to provide timely and reliably information to the college?

5.2.1 What managerial gaps are there in the constituent college?

This research question helped in understanding research problems, objectives, scope, risks and constraints involved. A survey was conducted to ascertain the level of dependency between top decision makers in the college and IT department in accessing, analyzing and reporting data. The survey also was meant to find out the level of competence in use of Ms Excel.

The results of the survey revealed that majority of the staff depend on IT professionals in producing reports. The survey also showed that a substantial number of respondents do not have easy access to the data and reports are not developed quickly as documented in section 4.4 of this report. At this Juncture, it is critical to mention that indeed there are managerial gaps in the college that need to be addressed.

5.2.2 What is the most valuable technology for the university users in accessing, analyzing and reporting data easily and quickly?

Literature survey was carried out to discover and comprehend the concepts, fundamentals architectures, and technologies involved in this project. Chapter 2 of this report strived to make certain on the above aspects with more emphasis on data warehousing, a technology that is applied within Business Intelligence (BI) systems. As stated earlier, a BI system is a systematic use of information about your company and its business environment to analyze, report, predict and manage business performance. Section 2.4 of this report considered data warehouse as one of the most valuable technology for the purpose of this project. Section 4.2 delineated OLAP Reporting tool based on Excel as an example of an outstanding Business Intelligence (BI) system to instigate the technology in question.

5.2.3 How will the technology be implemented to narrow the managerial gaps in the college?

The data warehousing technology under discussion consolidated essential historical data from the underlying college disparate databases. In section 3.3, source data was identified and organized into main subjects such as Finance, Enrolment and Performance. Each of this

subject area constituted a multidimensional-cube in the resultant data warehouse database. It is worthy to stress again that this project is designed to implement a single multidimensional-cube database as mentioned in the project scope (section 1.6) of the report.

This study considered and applied Star Schema dimensional modeling techniques to design an integrated data structure required in a data warehouse. Since data was from disparate sources, transformation process was carried out which involved conversion, summarization and filtering of data. Transformed data was then loaded into a data warehouse using Database Transformation Services (DTS) tool of Microsoft SQL 2000 as shown in section 4.1 implementing data warehouse.

5.2.4 How can the technology be optimized to provide timely and reliably information to the college?

OLAP Reporting tool based on Excel gives Excel users and more specifically decision makers significant power and flexibility to access, analyze and reports data. Such OLAP program was well illustrated in section 4.2 while its effectiveness and efficiency was demonstrated in section 4.3

5.3 Research Achievements

The survey conducted turned out to be successful in light of this project. A sample size of 97 members of staff was selected from the entire population of about 430 and issued with questionnaires. The number of questionnaires that were filled and returned by respondents from all college sectors under consideration were 74. This represented 76.3% of the sample size. Moreover, the results of the survey reflected the current position in the college as far as data analysis and reporting was concerned.

This project placed strong emphasis on the data acquisition process which encompasses sourcing, cleansing, transforming, and aggregating data using Database Transformation Services (DTS) tool of Microsoft SQL 2000. Through the cleansing process, the study enhanced data quality by ensuring data accuracy, type, and consistency, as well as eliminating duplicate records

The study has successfully implemented a data warehouse database composed of a single multidimensional cube based on subject area called Finance. This database brought together data from multiple sources providing principals and decision makers of the college greater insight into the college financial performance.

The research successfully implemented Business Intelligence (BI) program based on Excel that turned ordinary spreadsheets into a flexible, powerful, and inexpensive BI system. OLAP Reporting tool based on Excel was used as BI software. The bottom line is that Excel-friendly OLAP programs gives Excel users significant power and flexibility, while eliminating many problems associated with writing values to cells.

5.4 Summary of Research Findings

Most typical reports take a least a day to be prepared. Majority of the respondents are of the same opinion that currently, typical reports takes at least a day to prepare

The users of the reports hardly prepare them without the help of IT people. Greater part of users is of same mind, despite the fact that some argue that they can create a report on their own

Many decision makers understand what the data means, but they don't have easy or dynamic access to it. Half of the respondents strongly concur, as some say they are satisfied the way they interact with data.

Some of the reports prepared either contain errors or doctored to suite its purpose. Almost the same number of respondents is for or against the claim that reports created are with much less possibility of errors.

Many users are accustomed to working with many versions of truth. A third of the staff disagrees and claims that similar reports are generated several months later against similar data. However, a small number of respondents claim that they have been creating dissimilar

reports due to frequent change of data and therefore sometimes they are forced to maintain a spreadsheet database with each month's data.

The quality of the reports produced is not up to standard. Only 21.7% of the respondents feel satisfied with the quality of data analysis and reporting. 40.5% rate it poorly and feel very dissatisfied

Majority of university staff are competent excel users although more than half of the them have not used Ms Excel as an outstanding Business Intelligence (BI) tool while 8.11% believe they have done so

5.5 Recommendations

The major source of income to the university college comes from diploma courses as exemplified in section 4.3.2. The college management team therefore, should explore a larger percentage of students who do not qualify for degree programs but have the potential to continue pursuing after completing secondary education. This study pointed out in section 1.1 that only 81,000 of 250,000 qualify to join public or private universities. This leaves a bigger percentage of students the college must continue to tap despite being upgraded to a university status.

The college management must also appreciate that its potential market or target population as stated above, do have the ability to pursue non technical courses. Therefore, University College courses should be restructured so that the college may offer courses that can generate more revenue. This has been clearly demonstrated in section 4.3.4 where highest income generating course is mass communication at diploma level. It is also important to note that another course in the same category with mass communication at the degree level was the only course that generated more income in 2011 while all other degree courses dropped at the same period.

The college should implement a data warehouse using existing resources to address managerial gaps that hinder decision making process. This study has shown that existing resources or equipments, the prevailing environment, and manpower the college has can meet 80% of the required resources to realize data warehousing system. For that reason, the issue

of cost should not arise instead complexity involved in implementing a data warehouse may be looked into and be varied.

The principals and senior decision makers should be provided with an integrated, user-friendly interface that will dynamically allow them to access required data, analyze it and transform results of the analysis into a report with minimal intervention from IT department.

5.6 Limitations and Constraints

The depth and breadth of this investigation study which includes an extensive literature review initially put a restriction on the inclusion of the OLAP technology and data mining within the main report. However, after a thorough review, the main content of the report concentrated on problem identification, modeling design and data warehouse implementation. As a result, less emphasis has been put on business analysis of the user requirements and data mining using data mining tools.

5.7 Areas for future work

This investigative study can be further extended in several ways. First, a larger number of data warehouse schemata such as the fact constellation schema, multi star schema, starER schema and so on can be included for investigation. Secondly, the consistency of the data warehouse system performance can be further investigated using larger datasets for the dimension tables and with different aggregation levels. Finally, an interesting extension is to include the first two expansions in a study and to derive guidelines from the findings to support the data warehouse designer and practitioner in a more efficacious and comprehensive manner.

5.8 Summary

Data warehouses collect, consolidate, organize, and summarize data so it can be used for business decisions. Dimensional modeling, the foundation of data warehouse design, is not an arcane art or science; it is a mature methodology that organizes data in a straightforward, simple, and intuitive representation of the way business decision makers want to view and analyze their data.

The key to data warehousing is data design. The business users know what data they need and how they want to use it. Focus on the users, determine what data is needed, locate sources for the data, and organize the data in a dimensional model that represents the business needs.

The remaining tasks flow naturally from a well-designed model—extracting, transforming, and loading the data into the data warehouse, creating the OLAP and data mining analytical applications, developing or acquiring end-user tools, deploying the system, and tuning the system design as users gain experience.

Microsoft SQL Server 2000 provides a wide range of powerful and easy to use tools you can use to create a data warehouse and analyze the data it contains. The ability to design and create data warehouses is no longer isolated to experts working with primitive implements.

REFERENCES

Ackoff, RL 1989, 'From Data to Wisdom', *Journal of Applied Systems Analysis*, vol. 16, pp. 3-6

Akintola KG, Adetunmbi, AO, Adeola, OS, 'Building Data warehousing and Data Mining From Course Management Systems': *A case of Federal University of Technology* (FUTA) Course Management Information Systems. (ITePED 2011)
www.ncs.org.ng

Anthony, R. A 1965, *Planning and Control Systems: A Framework for Analysis* (Harvard University Graduate School of Business Administration, Cambridge, M. A)

Ballard, C, Herreman, D, Schau,D, Bell, R, Kim, E, Valencic, A, 'Data Modeling Techniques for Data Warehousing'. International Technical Support Organization, SG24-2238-00
http://www.redbooks.ibm.com

Barry, D 1997, *Data Warehouse from Architecture to Implementation*, Addison-Wesley.

Curtis, G 1995, *Business Information Systems: Analysis, Design and Practice*, Second Edition, Addison-Wesley.

Dell'aquila, C, Di Tria, F, Lefons, E, Tangorra, F, 'An Academic Data Warehouse'

Dell'Aquila, C, Ditria, D, Lefons, E, Tangorra, F 2007, 'An Academic Data Warehouse'. *Proceedings of th 7th WSEAS International conference on Applied Informatics and Conications*, Attens, Greece, August 24-26

Furlow, G, 'Data Warehousing, the Case for Building a Data Warehouse', 1520-9202, July |August 2001 IT Pro.

Gallegos, F 1999, 'Data Warehousing: A Strategic Approach'. *Information Strategy: The Executive's Journal,* Vol. 16, No. 1, Fall 1999, pp. 41-47

Gorry, GA & Scott Morton, MS 1971, 'A Framework for Management Information Systems', Sloan Management Review, 13(1), 1-22

Inmon, WH, 'Building the Data Warehouse: Getting Started'

Inmon, W 1996, *Building the Data Warehouse*, John Wiley & Sons, New York;

Jarke, M, Lenzerini, M, Vassiliou, & Vaassiliadis 2003, *Fundamentals of Data Warehouses*, 2nd Edition, Springer-Verlag.

Kimball, R 1996, *The Data Warehouse toolkit: Practical Techniques for Building Dimensional*

Lechtenbörger, J & Vossen, G 2003, 'Multidimensional normal forms for data warehouse design' (Information Systems 28 (5), 415-434)

Microsoft Populating a Data warehouse with Microsoft SQL Server 2000 Data Transformation Services, 2092A

Pendse, N 2005b, 'Market share analysis', OLAP report Available from: http://www.olapreport.com/market.htm

Pendse, N 2005a 'OLAP architecture'. OLAP report Available from: http://www.olapreport.com/Architectures.htm

Poe, V, Klauer, P, & Brobst, S 1998, *Building a Data Warehouse for Decision Support*, 2nd Edition, Prentice-Hall

Prasad, KVKK 2006, *Data Warehouse Development Tools. Covering Informatica, Cognos, Business Objects and DataStage with Case Studies.* Comprehensive and Practical Coverage of: Data warehousing concepts and design, 2006, 19-A

Scott Morton, MS 1971, *Management Decision Systems: Computer-Based Support for Decision Making*, Harvard University, Boston.

Simon, HA 1965, *Administrative Behavior*, (2nd Edition), New York: Free Press

Simon, HA 1977, *The New Science of Management Decision*, (rev. ed), Englewood Cliffs, NJ: Prentice Hall. (Original work publish 1960).

Velicanu, M, Matei, G, *Building a Data Warehouse Step by Step*, Informatica Economica, nr. 2(42)/2007

Weisensee, D, Mathews, E., McInnis A., ' Implementing Data Darehousing and Business Intelligence at MacMaster University Using the SAS Intelligence Value Chain' (Paper # 125-30)

Wierschem D., McMillen J., McBroom R. (2003), 'What Academia Can Gain from Building a Data Warehouse'. Used Effectively, Data Warehouses Can Be A Significant Component Of Strategic Decision Making On Campus, No. 1, Educause Quarterly 41.

Appendix

QUESTIONNAIRE

UNIVERSITY DATA WAREHOUSE SURVEY

Please take a few moments to complete our University Data Warehouse Survey. Your feedback is important to us. Your responses will be kept confidential and will not be used for any purpose other than research. This survey will take approximately 10 minutes to complete.

Section A: Demographic Information

This section will be used to find out demographic information that will help us cater our respondent's needs.

A.1 Gender

☐ Male ☐ Female

A.2 Age

☐ 18-34 ☐ 35-44 ☐ 45-54 ☐ 55 or Above ☐ Prefer not to answer

A.3 Employment

☐ Part-Time ☐ Permanent ☐ Contract ☐ Prefer not to answer

A.4 Salary

☐ 49.999 or Less ☐ 50.000 – 74.999 ☐ 75.000 – 99.999

☐ 100. 000 – 119 999 ☐ 120.000 or Above ☐ Prefer not to answer

A .5 Education

☐ Masters and above ☐ Undergraduate ☐ Diploma

☐ Form 4 or Less ☐ Prefer not to answer

Section B: Data Analysis and Reporting

This section contains questions that will accurately measure producers of reports and duration it takes to create them. Your feedback will help us realize the level of dependency, existence of dynamic data access, quality, and accuracy of reports produced. It will also measure level of satisfaction in analyzing and reporting data.

As you answer the following questions, please think of your MOST RECENT college report produced in your department/section.

B.1 Who creates reports in your department/section?

☐ Technician ☐ Mvself ☐ Prefer not to answer

B.2 About how long does it take to create a report?

☐ More than a day ☐ One dav ☐ 4 hours

☐ Less than 2 hours ☐ Prefer not to answer

Please tell us how well you agree with the following statements.

		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
		5	4	3	2	1
B.3	Reports are produced with the help of technician					
B.4	I cannot dynamically access all required data					
B.5	Dissimilar reports produced later					
B.6	Reports often contain errors due to source data					
B.7	Reports are not developed quickly and easily					
B.8	Report users don't have direct access to operational data(base)					

B.9 What is your overall satisfaction level of analyzing and reporting data?

☐ Satisfie ☐ Neutral ☐ Dissatisfie ☐ Prefer not to

Section C: Ms Excel Users

This section will be used to find out whether you have used Ms Excel before and for how long. The section will also enable know whether you have used Ms Excel before as an outstanding Business Intelligence tool or if it can be used for that purpose.

D.1 Have used Ms Excel before? ☐ Yes ☐ No

(If the answer is No, please skip this section.)

D.2 How long have you used Ms Excel?

☐ Less than a vear ☐ 1-4 vears ☐ 5- 9 ☐ Over 10 vears

D.3 What type of user are you in using Ms Excel?

☐ Expert ☐ Experienced ☐ Causal ☐ Not Sure

D.4 In the past, have you used Ms Excel as an outstanding Business Intelligence (BI) tool?

☐ Yes ☐ No

D.5 Ms Excel can become a key component of a powerful and inexpensive BI tool?

☐ Definitelv ☐ Probablv ☐ Not Sure ☐ Probablv Not ☐ Definitelv

Thanks for your time and valuable input!