Sudan University of Science and Technology
College of Graduate Studies

Thesis for the Doctor of Philosophy
Degree in Computer Science

Analysis of Factors Affecting Higher Education Learning Process using a Bussines Intelligence Model

تحليل العوامل المؤثرة على العمليه التعليمية في التعليم العالي باستخدام نموذج ذكاء أعمال

By:

Wafa Ali Mohammed Abdelrhman

Supervised By:
Prof. Fanny Klett

December-2020

قال تعالى:

﴿ السَّمَاوَاتِ وَمَا فِي الْأَرْضِ ۗ وَإِن تُبْدُوا مَا فِي أَنفُسِكُمْ أَوْ
ۖ فَيَغْفِرُ لِمَن يَشَاءُ وَيُعَذِّبُ مَن يَشَاءُ ۗ وَاللَّهُ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ (284)
لَيْهِ مِن رَّبِّهِ وَالْمُؤْمِنُونَ ۚ كُلٌّ آمَنَ بِاللَّهِ وَمَلَائِكَتِهِ وَكُتُبِهِ وَرُسُلِهِ
مِن رُّسُلِهِ ۚ وَقَالُوا سَمِعْنَا وَأَطَعْنَا ۖ غُفْرَانَكَ رَبَّنَا وَإِلَيْكَ الْمَصِيرُ (285)
نَفْسًا إِلَّا وُسْعَهَا ۚ لَهَا مَا كَسَبَتْ وَعَلَيْهَا مَا اكْتَسَبَتْ ۗ رَبَّنَا لَا تُؤَاخِذْنَا
خْطَأْنَا ۚ رَبَّنَا وَلَا تَحْمِلْ عَلَيْنَا إِصْرًا كَمَا حَمَلْتَهُ عَلَى الَّذِينَ مِن قَبْلِنَا
رَبَّنَا وَلَا تُحَمِّلْنَا مَا لَا طَاقَةَ لَنَا بِهِ ۖ وَاعْفُ عَنَّا وَاغْفِرْ لَنَا وَارْحَمْنَا ۚ أَنتَ
مَوْلَانَا فَانصُرْنَا عَلَى الْقَوْمِ الْكَافِرِينَ ﴾

[البقرة: 284، 286].

Dedication

I dedicate this dissertation to my parents, for their endless support and encouragement. To my husband *Mohsen*who has been a constant source of support and encouragement during the challenges of research. To my kids *Mounia*, *Menn*, *Yomna* and *Mohmmed*. To my brothers and sisters, and all their children for their support.

Acknowledgement

First I would like to thank my research advisor *Prof.Fanny Klett* for her guidance and support through this study. Also thank to *Prof. Izzeldin Osman*who is father of Computer Science PhD program.

Many thanks are extended to all SUST staff, and to my colleagues in PhD program especially *Izzeldin Ali*for his support and encouragement.

I am grateful to my family for their kindness, love and support. They are always my strong backing.

Also many Thank to everyone who helped me in obtaining the necessary data and informationthrow this research journey.

Abstract

Higher Education Institutions (HEIs) today face many challenges such as the local and international expansion and the widespread utilization of Information and Communication Technologies (ICT) that facilitate global competition. To deal with these and with the changes in labor market, HEIs administrators and planners need information to improve their services, satisfy customers (students) and gain competitive advantage. Learning is a process that includes interactions among students, educators and courses under the umbrella of HEIs. In fact, it is the core and most important service provided by them and it is surrounded by many factors. Improving the learning process is multifaceted as there are numerous factors and stakeholders involved. The aim of this study is to present a Business Intelligence (BI) solution and investigate its value in streamlining the learning process at HEIs. The proposed solution avails information and analytical capabilities to HEIs decision makers. A mix of research methods were used to validate the solution, including content analysis to identify and categorize the factors involved in the learning process, along with a case study implementation of the proposed BI solution. Data was collected from different departments at the Sudan University of Science and Technology (SUST). The proposed model consist of two environment: data warehouse environment and analytical environment. The findings obtained from the model have demonstrated the efficiency and benefit of using BI tools to analyze data and obtain valuable decision-making information. The implemented model is very robust, and additional data will lead to more information about factors, and the relationship between factors and learning process.

المستخلص

تواجه مؤسسات التعليم العاليااليوم العديد من التحديات مثل التوسع المحلي والدولي والاستخدام الواسع لتقنيات المعلومات والاتصالاتالتي تسهل المنافسة العالمية. للتعامل مع هذه ومع التغيرات في سوق العمل ، يحتاج مديرو ومخططي مؤسسات التعليم العالي إلى معلومات لتحسين خدماتهم وإرضاء العملاء (الطلاب) واكتساب ميزة تنافسية. التعلم هو عملية تتضمن تفاعلات بين الطلاب والمعلمين ودورات تحت مظلة مؤسسات التعليم العالي. في الواقع ، إنها الخدمة الأساسية والأكثر أهمية التي يقدمونها وهي محاطة بالعديد من العوامل. إن تحسين عملية التعلم متعدد الأوجه حيث يوجد العديد من العوامل وأصحاب المصلحة المعنيين. الهدف من هذه الدراسة هو تقديم حل ذكاء الأعمالوالتحقيق في قيمته في تبسيط عملية التعلم في مؤسسات التعليم العالي. يوفر الحل المقترح المعلومات والقدرات التحليلية لصناع القرار في مؤسسات التعليم العالي. تم استخدام مزيج من طرق البحث للتحقق من صحة الحل ، بما في ذلك تحليل المحتوى لتحديد وتصنيف العوامل المتضمنة في عملية التعلم ، إلى جانب تنفيذ دراسة الحالة لحلذكاء الأعمالالمقترح. تم جمع البيانات من أقسام مختلفة في جامعة السودان للعلوم والتكنولوجيا. يتكون النموذج المقترح من بيئتين: بيئة مستودع البيانات وبيئة تحليلية.أظهرت النتائج التي تم الحصول عليها من النموذج كفاءة وفائدة استخدام أدوات ذكاء الأعمال لتحليل البيانات والحصول على معلومات قيمة لصنع القرار. النموذج الذي تم تنفيذه قوي للغاية ، وستؤدي البيانات الإضافية إلى مزيد من المعلومات حول العوامل ، والعلاقة بين العوامل وعملية التعلم.

Table of Contents

Acronyms & Abbreviations

Acronym	Description
HE	Higher Education
HEI	Higher Education Institution
ICT	Information & Communication Technology
BI	Business Intelligence
DSS	Decision Support System
ETL	Extract, Transform and Load
DW	Data Warehouse
EDM	Educational Data Mining
CPM	Corporate Performance Management
BPM	Business Performance Management
EIS	Executive Information System
OLAP	Online Analytical Processing
SRM	Student Relationship Management
CRM	Customer Relationship Management
EI	Educational Intelligence
TOS	Talend Open Studio
KPI	Key Performance Indicator
BIRT	Business Intelligence and Reporting Tool
CIST	Computer Science and Information Technology
SUST	Sudan University of Science and Technology
CS	Computer Science
IS	Information System
NW	Network
SE	Software Engineering
WEKA	Waikato Environment for Knowledge Analysis
CSV	Comma Separator Values
HR	Human Resource
LMS	Learning Management System

Tables

Figures

Appendices

Publications

Paper One Wafa Ali Mohammed *"Investigating Factors Affecting the Learning Process in Higher Education - A Case Study"* International Journal of Computer Trends and Technology 68.6 (2020):10-17.

Paper Two Wafa Ali Mohammed, Izzeldin A. Elhassan *"A Business Intelligence Solution for Improving the Higher Educational Learning Process"* International Journal of Computer Trends and Technology 68.12 (2020):15-20.

CHAPTER 1

1 Introduction

1.1. Background

Education is defined as a learning process for the individual to attain knowledge. It is an important process for individuals and society since it contributes to the progress and development of societies (Education 2000), (Teal 2011). No society can advance and become more civilized without education. The culture and values of any society depend on the education of individuals as cultural values are passed on through educational institutions. Generally, the formal educational process is composed of the following stages, kindergarten, primary, middle, secondary and the higher education (HE) HE stage. Higher Education Institutions (HEIs) deal with post-secondary level of education, and this includes undergraduate and postgraduate colleges, universities and centers.

The core functions of HEIs are: education, research and contributing to the society. In addition, they play a vital role of supplier to the labor market, as they provide highly qualified human resources. Learning is the most important process in HEIs, and it involves three main stakeholders, students, educators, and institutions.

HEIs currently face a variety of challenges, as globalization has unlatched international competition in addition to the local one. Moreover, the surge in the number of students without financial support from governments, and the expansion in HEIs both locally and globally. Confronting these challenges requires strategic and tactical planning by HEIs as they must maintain their existing students and stay competitive in recruitment by providing a service that satisfies their customers (students). HEIs must also respond to and stay in-line with the continuous changes in the labor market. Focusing on learning and how this service can be provided in a modern, efficient, and satisfactory to students will make HEIs more competitive.

Innovations in Information and Communication Technology (ICT) have enabled HEIs to better respond to their growing challenges and achieve their objectives. Currently, almost all HEIs utilize ICT in different capacities to support students' operations, HE management, administrative operations, and planning.Using of ICT lead to generate high volume of data but the managerial decisions are rarely taken based on it.Therefore, a high demand for information technologies, methods and

tools arises that can beneficially transform data into information and knowledge. The Business Intelligence (BI) is an emerging solution for gaining information, nowadays it became an effective means for improving data analysis, decision making and trend investigation. BI gives organizations a comprehensive and integrated view of their business operation and potential, and this facilitates the decision making processes(Falakmasir, Moaven et al. 2010).

1.2. Problem statements and its significance

Most HEIs in Sudan implement the traditional educational system (face to face classroom). The data collected at admission of newstudents, during the learning process and for management issue.

However the large volume of data available about the learning process they have no more extracted information that can be used for improvement. To improve the learning process there is a need to know and investigate the factors that can affect this process in order to find weakness points and then decide about how to solve it. Thus there is high demand to analyze the data about learning process in order to get a knowledge that can help to decide about how the improvement can be done.

1.3. Hypothesis

In order to investigate the situation, relevant hypothesis will be setup in the following:

- The envisaged BI system can cover the collection, and observation of a big data set that allows for the analysis and extraction of valuable information and knowledge to be beneficially applied toward monitoring educational management goals, supporting decision making and investigating future trends in the educational development.
- The suggested solution will be able to increase the type and number of opportunities offered to more effectively deal with an organizational management in HEI, such as statistical reports, trend visualizations, real-time data analysis and visualization.
- The analysis of data can be proceed by appropriate tools selected on the basis of a thorough state-of-the-art review according to the thesis objectives.

1.4. Objectives and research questions

1.4.1. Objectives

The specific research objectives are:

- Investigate BI systems and their current applications, challenges and opportunities in HEIs.
- Identify and evaluate the factors that influence the learning process from the perspective of students.
- Design BI framework suitably analyze the data collected toward new opportunities for managing and improving HE processes.

1.4.2. Research questions

Q1: what are the main factors that affect the learning process?

Q2: How to effectively create the data models toward an easy-to-implement scalable BI framework leading to improved educational processes?

1.5. Thesis Structure

Chapter 2: Provides theoretical background of learning, learning process and factors that affect the learning process.

Chapter 3:Also background chapter about Business intelligence and its components.

Chapter 4: Covers the literature review that includes research about the use of BI and its components in HE.

Chapter 5: Methodology part and description of steps for designing model.

Chapter 6: discuss the proposed BI framework.

Chapter 7:provides a summary of the research, findings and contribution.

CHAPTER 2

2 Learning process in HEIs

2.1 Higher education institutions (HEI)

Education is the basis of the civilization of peoples and the reason for its progress and development, and different fields of science that may be obtained by the individual, but the academic education provided in schools and universities is the basis of different sciences and different world governments have been interested in education and developed laws and regulations and was able to divide it into specific stages of education, Kindergarten stage , Basic stag, Secondary stage and higher education stage.
Higher education is education provided by universities, colleges, and other institutions that award academic degrees. Higher education includes both the undergraduate (i.e., college) and the graduate (or postgraduate) levels(Allen 2017).
Higher education institutions are defined as academically qualified and applied educational institutions to teach all scientific, literary and professional disciplines to students according to the university system for each of them, in preparation for their involvement in the labor market
Higher education institutions have three functions these are: education, research and contributing to society. Research and education function are mostly tide together, education develop the human resources to do research and research make a high level of education possible. Contributions to society been demanded of higher education institutions(JICA 2004).

2.1.1 Importance of higher education

The most important reasons for people to seek to higher education institutions is career preparation this is importance for the individual and labor market because higher education provides human resources. Higher education helps an individual to get more job opportunities and often earn higher incomes than uneducated individuals.Higher levels of educational achievement lead to higher earnings and lower unemployment (Card, 1999), along with lower crime, better health, and greater civic participation (Lochner, 2011). Higher education also add and improve some personal experiences as better communication this is done throw the advanced writing and speaking assignments; this trains individuals to express themselves clearly and communicate more effectively with

others. Critical thinking skills is also another experiences the person can achieve from higher education (Gulliver 2014).

2.1.2 Challenges

The environment surroundinghigher education sector today is a dynamic and complex environment (Kabakchieva 2015), globalization and transition from an industrial society to a knowledge society this increase the demand of higher education to produce the qualified human resources.Today's labor market requires highly skilled personnel at all levels to deal with rapid industrialization in rapidly changing environments(Chan 2016). To deal with market need and producing qualified alumni institutions needs to improve the learning process.

2.2 Learning

Learning occurs continuously throughout a person's lifetime. Learning is an essential activity and is even more important for youth because through it institutions are preparing and preparing them for their future role in the development of their societies and acquisition of success of life (Siming 2015, Fauzia 2014). There are many definitions for learning, Learning is a systematic relatively permanent change in behavior that occurs through experience.The *Oxford Companion to Philosophy* defines it as "the acquisition of a form of knowledge or ability through the use of experience." Another definition that learning is an activity that has purpose to make change of behavior, attitude, habitual, knowledge, skill, etc (Lachman 1997). Learning is done through providing and improving the learning process.

2.2.1 Learning process

The learning process is comprised of the interaction between institution, educator, student, environment and courses. The outputs from the learning process are the graduate students and their quality (Betoret and Tomás 2003)(Pavione, Avelino et al. 2016).

System is set of elements or componentsthat interact to accomplish goals(Stair and Reynolds 2013), from this definition and the above definition of learning process We can see learningas systemconsist of input, processing and output. The inputs are students, teachers, courses and any facility, the processing in the system is teaching process and the output from this system is graduate students who affect in community and labor

market Figure 2-1 give general view about learning process, the practice of learning process and the interaction between its elements (students, educator and course) is done in classroom context, real classroom in traditional education (face to face) or virtual classroom in e-learning, the general classroom affect the ongoing learning process. Classrooms are generally inside the HEI context, where all regulations, policies, environment and facilities provided by the institution affect the learning process. There is a strong bidirectional relation between the labor market and the HEI context, as labor market and community requirements affect the design of curriculum to produce graduates who in turn contribute to community and business development. The community is affected by the final output, because the development of community in all aspects such as, economic, political, social and cultural depends on the outputs of the learning process.

From the preceding informationand discussion, we conclude that the learning process is complex, interconnected and involves many factors and stakeholders. Identifying these factors and how they affect the learning process is an important issue for any improvement. Similarly, stakeholders play a role in the success or failure of the process, with students being the most important one. For two reasons, first they are at the heart of the process being the main customers for HEIs. Second, their success is the target and outcome of the whole process as graduates. Therefore, any effort to improve the learning process must take into account the view point of students. This includes investigating their attitude, expectation and perception toward various specific activities and characteristic features of the process, along with socioeconomic factors.

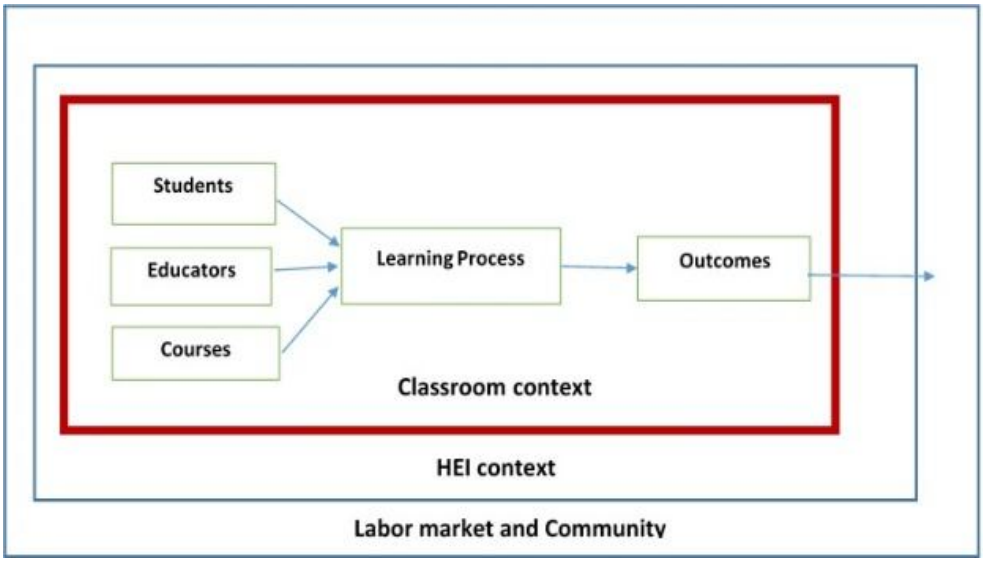


Figure 2-1 Learning Process

2.2.2 Factors affecting the learning process

The process of learning is influenced by many factors related to student, family, class, colleagues, educators, courses, teaching methods, technology, HEI policy, regulations, aids and the general environment. These factors can be grouped into six mainfactor groups or dimensions, namely: environment, student, teaching, academic advising, HEI, and ICT.

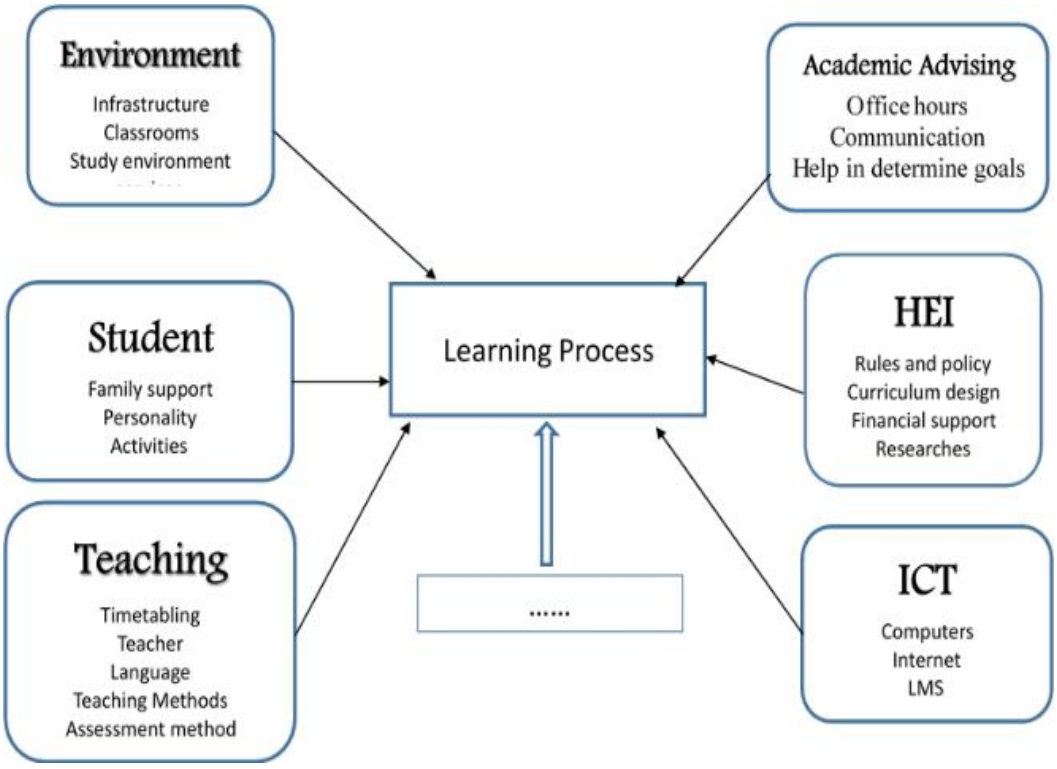


Figure 2-2 Factors Affecting Learning Process

Figure 2-2 presents these six dimensions with arrays of their comprising factors or characteristics. These arrays are dynamic and can expand or reduce in size over time.

2.2.2.1 Environment

Learning environment refers to the space allocated for classrooms, science labs, open spaces and offices. Also learning environment refers to diverse physical locations, contexts and culture in which students learning (Amirul, Ahmad et al. 2013). From this definition we can see that environment is about different physical aspects that affect the learning process inside classroom (light, sound, air condition, seats …) and outside classroom. The categorization of this aspects can be as follow:

- Infrastructure environment

  Infrastructure as road, building, building arrangement, garden, parking…etc. all this are important because if the roads inside university are well paved and the building adjacent this lead to easily movement and time preventing and this satisfy students.

- Classroom environment.

  One of the areas that make an obvious impact on student success is the physical environment of theclassroom, because most of student's time is spending inside classroom (Amirul, Ahmad et al. 2013).There is relationship between the physical structure of classroom as light, sound, seats, size, temperature and the psychological factors as behaviors and interaction in classroom(Puteh, Che Ahmad et al. 2015).

  The existence of labs, university libraries, cafeterias and dedicated places to practice activities are necessary elements of the good educational process, which are prerequisites for achieving educational goals.

  The HEI environment is one of the most important factorsthat affect the behavior of students and their achievement and their direction towards study.

2.2.2.2 Student:

The student plays an important role in the success of the learning process. There are many student characteristics that influence the process. First, is motivation or desire, and this the driving force behind student learning and success(Aslam, Younis et al. 2012). Second, is family environment and support. This is characteristic is reflected in students' relation to their parents, their communication with them, and their family income. Students who live in stable family environment and find good support, do well in their studies. Third, the relationship and communication with colleagues.

2.2.2.3 Teaching

This factor contains some important sub-factor:

- *Teacher and teaching aids*

Teacher is most important elements of the educational system, the roles of teacher as a faculty member differed, and he became facilitator of the learning process, leading learning from one educational step to another. Teacher play an extremely important role

in guiding students to achieve their academic goal (Aslam, Ali et al. 2012). Teacher who has solid and up-to-date knowledge and encourage contact with students inside and outside classroom is one of the basic elements that influence the learning process. According to (Pavione, Avelino et al. 2016) there aretechnical and personal requirements in the teacher dimension, such as: appropriate verbal instructions, which can be interpreted as the method of teaching; information to students about their progress; good relationship with the student; and attitude towards the subject taught.

Teaching aids is another element that can affect learning process, teaching aids are those materials that help the teachers explain better knowledge and help the students understand. Teaching aids play important role in learning process, it motivate the students so that they can learn better also the teacher clarify the information easily in addition to that the teaching aids make the classroom live and active. They are many types of teaching aids visual aids, audio aids and audio - visual aids(Nikky 2010).

- *Courses*

Sequence of courses, contents and how it meet students need, curriculum design and if it relate to labor market all this have great affection of learning process. The content of course should meet students need (Pavione, Avelino et al. 2016)and objective of course must be cleared and sequence of contents is important. Related to contents is how the information are display or presents to students the teacher must take in their account the different learning styles in classroom. Learning style is about how student absorb information and analyze it, learning styles can be categorized in three modes: first one is visual students who belong to this modes like to see the what the teacher mean and make decisions based on how things look. The second mod is Auditory (also called Aural or Audio) this group prefers to hear, the third mode is kinesthetic students in this mode like to experience what the teacher talk about. Some students are multimodal this group of students can use more than one mode(Abante, Almendral et al. 2014). An important thing related with course is assessment method, assessment methods are the strategies, techniques, tools and instruments for collecting information to determine the extent to which students demonstrate desired learning outcomes. Assessment method allows teachers to ensure students learn what they need to know in order to meet courses learning objectives, using multiple assessment methods will lead to complete view about student’s achievement.

- *Teaching methods*

Teaching method is the way of teacher deliver the subjects to students by using certain methods corresponding to characteristics of the students that were encountered. There are different kind of teaching methods as lecture, discussion, question and answer and group work. Choose more appropriate method for teaching lead to increase learning achievement(Munawaroh, 2017). There is a relationship between the teaching method and the classroom environment

- *Timetable of lectures*

Timetable setting is important issue, because it affects learning process.The timetable affects the educational process, if the table is compressed and does not contain any space between the lectures, it negatively affects the student. The student needs time between lectures to practice other activities such as eating, drinking and praying, and also needs a rest of mind so that he can receive new information.

2.2.2.4 Academic advice

Advising is a process in which advisor and advisee enter a dynamic relationship respectful of the student's concerns. Ideally, the advisor serves as teacher and guide in an interactive partnership aimed at enhancing the student's self-awareness and fulfillment (Cook 1999). Advising process can help students to identify personal strengths and interests related to their educational and career goals(Young-Jones, Burt et al. 2013). Academic advising is important because it is a central key to any student's success in college. The functions of adviser are to monitor students' academic progress, Give students the ability to navigate the university policies and procedures better and Assist students in their problems. According to these functions' continuous interactions between students and academic advisor is important and help in increasing student retention (Lau 2003) and this plays vital role in improving learning process.

2.2.2.5 HEI

The HEI itself plays a key role in the educational process, providing the necessary environment, designing curricula related to the labor market, establishing clear and understandable laws and regulations for students, providing financial assistance to poor students, providing incentives for superior students, and developing research projects and reflecting their results on the educational process. All this affects directly on quality and efficiency of the process.

2.2.2.6 ICT

It is obvious that ICT utilization and its necessary support have strong influence on all stakeholders of the learning process. They do so in a number of ways, such as, facilitating educators' access to information, the preparation of educational materials and the presentation of information to students. ICT provides the possibility of sharing information between educators and students, and students among themselves. It also enables opportunities for discussion and exchangeof ideas through the use of academic sites, social media and collaboration work techniques(Kumar). Furthermore, the availability of adequate numbers of computers,high speed internet, Wi-Fi, secure data processing and support are extremely important for students and facilitates their learning.

2.2.3 Summary

From discussion of the learning process in above sections and the factors related to this process we can conclude that the learning process is vital process and to cope with labor market needs and changes there is always a need to improve this process. The improvement is done throw the studying the factors and trying to find the negative and weakness points and try to solve and improve this factors. The list of factors is not an endless list its scalable list as world around the learning process is change this list can increased over time.

CHAPTER 3

3 Business intelligence as a resolving technology approach

3.1 Decision making

The business environment is change in quick manner, organizations in such environment require to take more decisions to cope with this change. Making decisions require more data, information and knowledge.The decision making is process of choosing among two or more existing alternative. According to Simon (Simon 1997) the decision making process consist of three phases: intelligence, design and choice. The first phase, the so called intelligence activity, is about searching the environment for conditions that requires decisions. The second phase includes inventing, developing, and analyzing possible responses to the identified conditions and is called the design activity. The third phase is called the choice activity and involves a selection of particular responses to the noticed conditions(Asemi, Safari et al. 2011), Figure 3-1 present the phases.

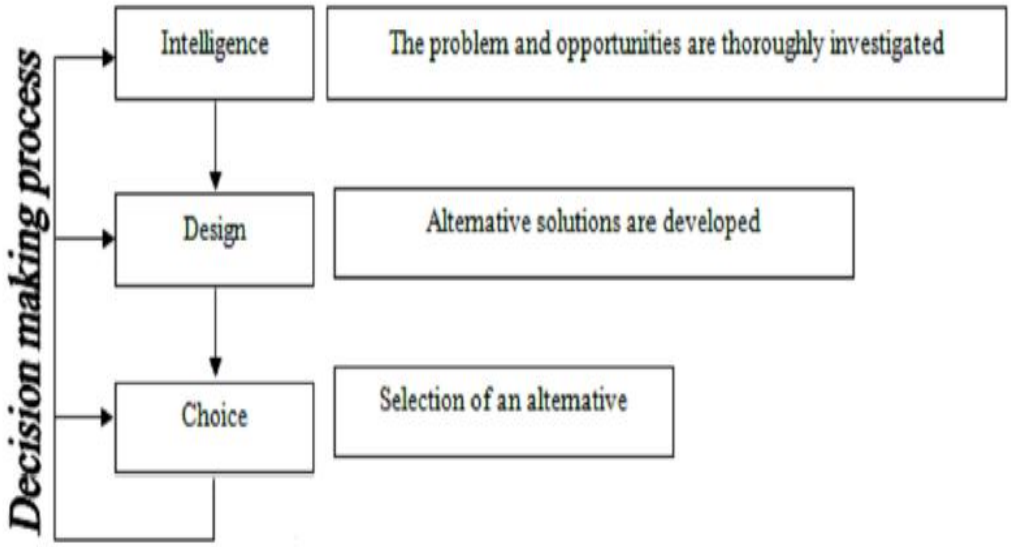


Figure 3-1The Decision-Making Process *(Asemi, Safari et al. 2011)*

The model of Simon is expanded by Huber whoadded two phases into Simon'smodel, fourth phase called implementation phase here the solution is put into effect. The last phase monitoring phase. In this phase,decision makers evaluate the implementation to determine whether the anticipated resultswere achieved and to modify the process based on new information(Asemi, Safari et al. 2011).

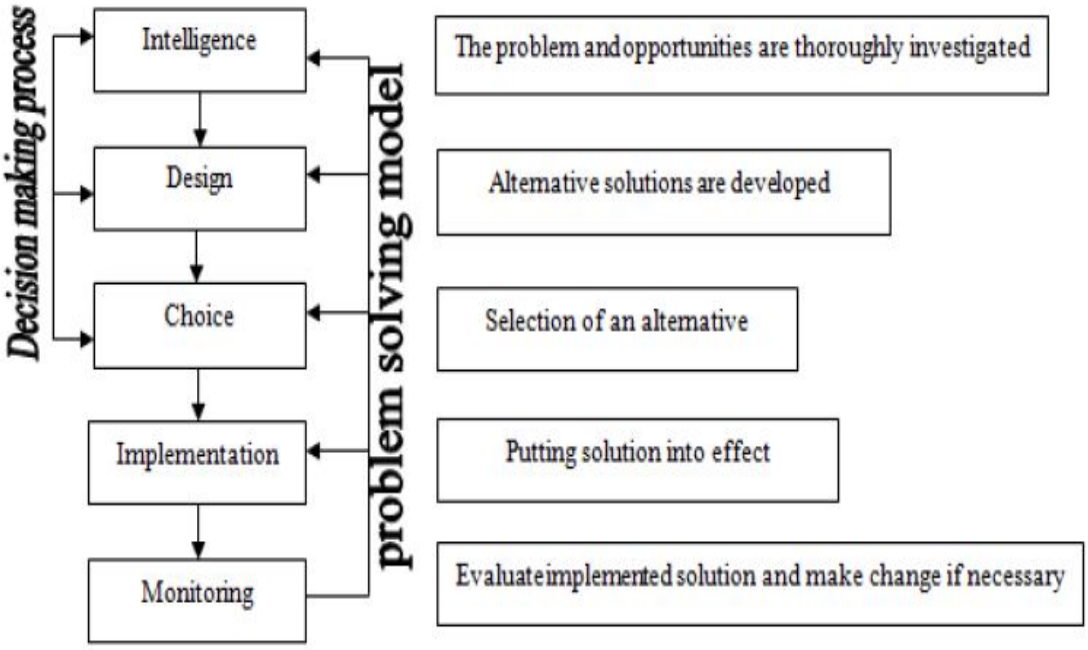


Figure 3-2 Problem solving model*(Asemi, Safari et al. 2011)*

There are two type of decision:

Programmed decisions: are those that are repeated over time and for which an existing set of rules can be developed to guide the process. These decisions might simple, or they could be fairly complex, but the criteria that go into making the decision are all known or can at least be estimated with a reasonable degree of accuracy.

Non-programmed decisions: are novel, unstructured decisions that are generally based on criteria that are not well-defined. With non-programmed decisions, information is more likely to be ambiguous or incomplete, and the decision maker may need to exercise some thoughtful judgment and creative thinking to reach a good solution. These are also sometimes referred to as non-routine decisions or as high-involvement decisions because they require greater involvement and thought on the part of the decision maker(Pomerol and Adam 2004).

For all types of decision there is a need of information, Therefor the importance of supporting institutions with robust and effective information systems to provide right information at the right time to decision makers has emerged. Especially in the case of non-programmed decision(Turban 2011).

3.2 Business intelligence (BI)

Business environment toady becomes more complex since there is constant change, organizations work in such environment need to respond quickly to change. Such activities require to take decisions. Decision making require amount of data, information and knowledge. Using computerized support system as BI help managers to get right information in right time to take decisions quickly(Turban 2011).

3.2.1 Definition

The original BI is evolved from DecisionSupport System (DSS)which is computer based support system for management decision maker(Watson 2009). With the development of DSS the term Business intelligence is introduced by Howard Dresner in 1989 who defined business intelligence as an umbrella term to describe "concepts and methods to improve business decision making by using fact-based support systems "(Power 2007). Another definition by Solomon and Paul is BI systems combine data gathering, data storage and knowledge management with analytical tools to present complex internal and competitive information to planner and decision makers(Solomon and Paul 2004).Also Watson defines business intelligence as a professional discipline, prescribes organizational and technological interventions aimed at providing timely, accurate information to the right people, at the right times, to enable and improve business decision-making(Watson 2009).

Another definition is BI refers to computer-based techniques used in spotting, digging-out and analyzing business data(Hema and Malik 2010).

Saeed_rouhani,et al defined BI as business intelligence is a set of abilities, tools, techniques and solutions that help mangers to understand business situation(Rouhani, Asgari et al. 2012).

Acording to Rouhani et alBusiness intelligence is defined as the ability for an organization to take all its process and capabilities and then convert these into knowledge, ultimately getting right information for the right people, at the right time through the right channel(Rouhani, Asgari et al. 2012).

IBM def. BI is the gathering, managing, analyzing and sharing of information in order to gain insights that can be used to make better decisions. BI turns information into intelligence, intelligence into knowledge, and knowledge into business wisdom.

There is no universal definition but the common goal of using business intelligence is to improve business decision making process and the benefits of using business intelligence are saving time and reducing the cost of IT infrastructure (Kelly 2005).

3.2.2 Components of Business intelligence

A Business Intelligence (BI) solution can include the following components, figure 3-3 present the components: data warehouse, business analytics, business performance management and the user interface. The next sections discuss this component in details as it's the core of suggestion framework.

3.2.2.1 Data warehouse (DW):

DW is a most important component of business intelligence. It's a single organizational repository of enterprise wide data across many or all lines of business and subject areas. Contains massive and integrated data, Represents the complete organizational view of information needed to run and understand the business.

In the late 1980's, IBM researchers Barry Devlin and Paul Murphy developed the concept of a data warehouse. Their idea was to provide a model for the flow of information from various operational databases to an environment that could be used for decision support(Bates 2015).

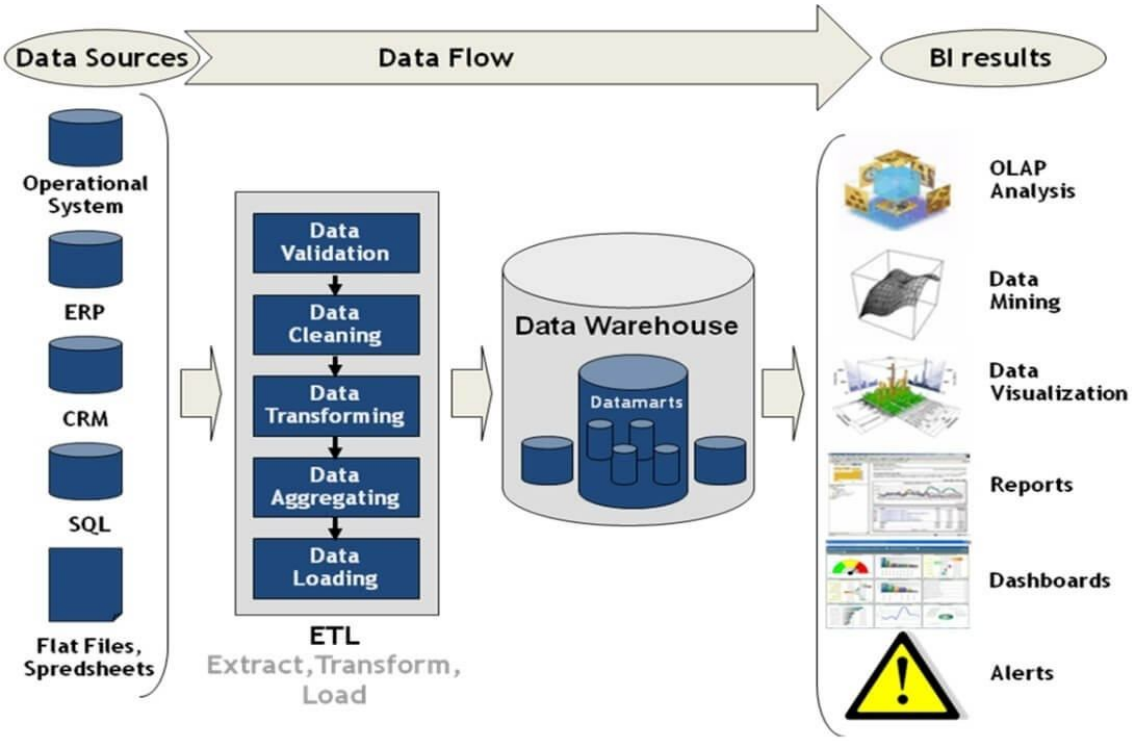


Figure 3-3 BI Components(Brown 2015)

According to Bill Inmon, who is a father of data warehouse concept provide the following definition: A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process(Inmon 2005).

- **Subject-oriented**, meaning that the data in the database is organized so that all the data elements relating to the same real-world event or object are linked together.
- **Time-variant**, meaning that the changes to the data in the database are tracked and recorded sc [illegible] ne.

Figure"": BI components ref.

- **Non-volatile**, meaning that data in the database is never over-written or deleted, once committed, the data is static, read-only, but retained for future reporting.
- **Integrated**, meaning that the database contains data from most or all of an organization's operational applications, and that this data is made consistent.

The data warehouse included only historical data that were organized and summarized, so end users could easily view or manipulate data and information. Today, some data warehouses include current data as well, so they can provide real-time decision support.

There are mainly four components of Data warehouse, figure 3-4:

- *Data source*

The data stored in data warehouse is derived from the operational systems that support the basic business processes of organization. Source data coming into the data warehouse may be grouped into four categories like production data, internal data, external data and archived data.

- *Staging area*

In this level operational data is cleaned and transformed to data warehouse. Three major functions performed in data staging are data extraction, data transformation and loading (ETL).

- *Data storage*

In this stage data is stored in the data warehouse.

- *Information delivery*

In this stage useful information is provided to users of data warehouse through various systems.

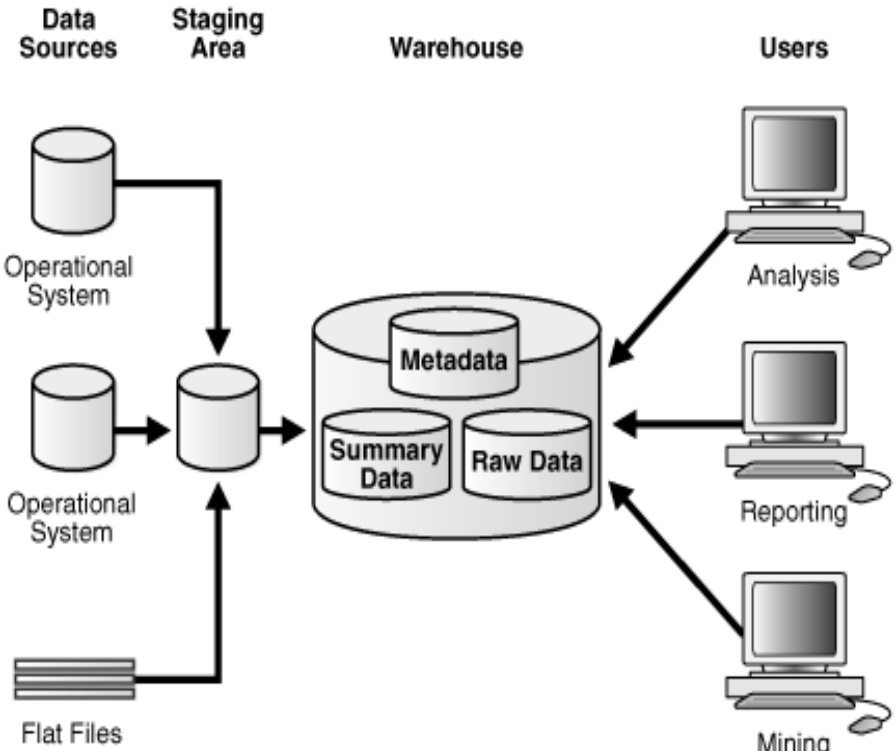


Figure 3-4DW Components(ORACLE®)

Data warehouse architecture

There are two general architecture used to design data warehouse one is the “Bill Inmon architecture” (Inmon 2005)and the other the “Ralph Kimball architecture”(Kimball and Ross 2011).

- *Inmon’s approach*

Inmon approach is a top-down he defines a data warehouse as a centralized repository for the entire enterprise. A data warehouse stores the “atomic” data at the lowest level of detail. The data marts are created only after the complete data warehouse has been created, figure 3-5.

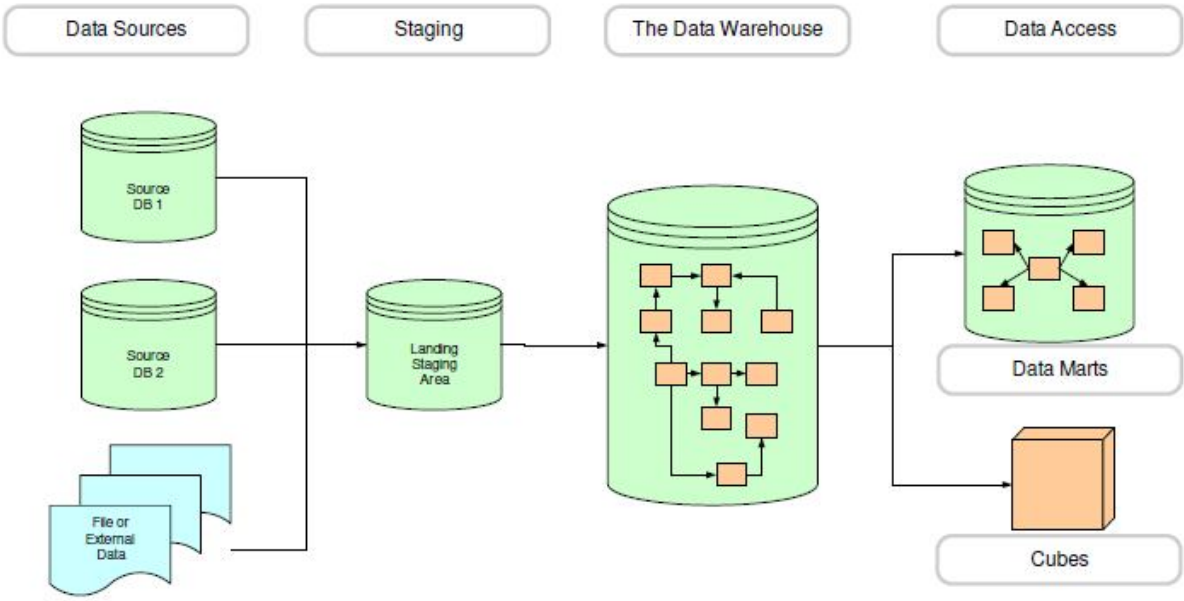


Figure 3-5 Inmon architecture(Abramson 2010)

- *Kimball’s approach*

Kimball’s approachis bottom-up approach he defines a data warehouse as “a copy of transaction data specifically structured for query and analysis”. In this approach the data mart is created firstly. Each data mart contains data specific to a business process, the data mart is build using dimensional model, figure 3-6.

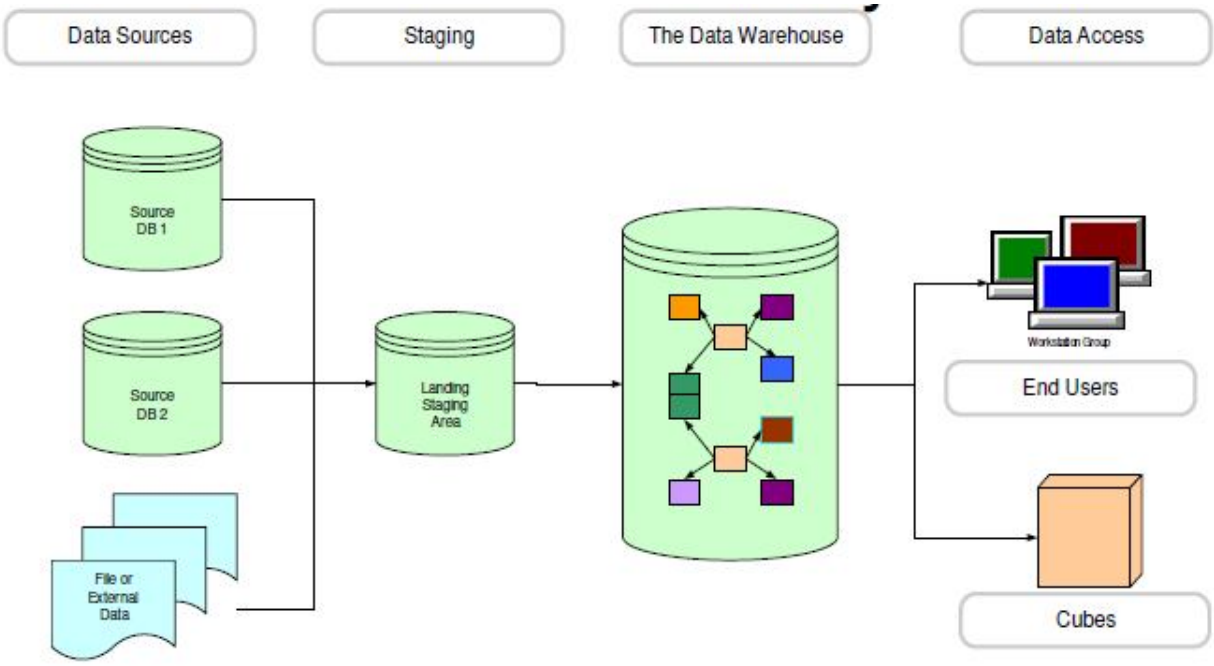


Figure 3-6Kimball'sarchitecture(Abramson 2010)

- Data Mart

Is specific, subject oriented, or departmental view of information from the organization, there are multiple data mart for organization. A data mart is more focused, generally smaller, selected facts and dimensions and integrated. Is built using dimensional modeling.

- Dimensional model

A dimensional model is a data structure technique optimized for Data warehousing tools. The concept of Dimensional Modelling was developed by Ralph Kimball and is comprised of "fact" and "dimension"tables. Facts are the measurements/metrics or facts from the organization business process. Dimension provides the context surrounding a two popular schemas to build the dimensional model:

- *Star schema*

The star schema architecture is easy to design. It is called a star schema because diagram resembles a star, with points radiating from a center. The center of the star consists of the fact table, and the points of the star is dimension tables. The fact tables in a star schema which is third normal form whereas dimensional tables are de-normalized figure 3-7.

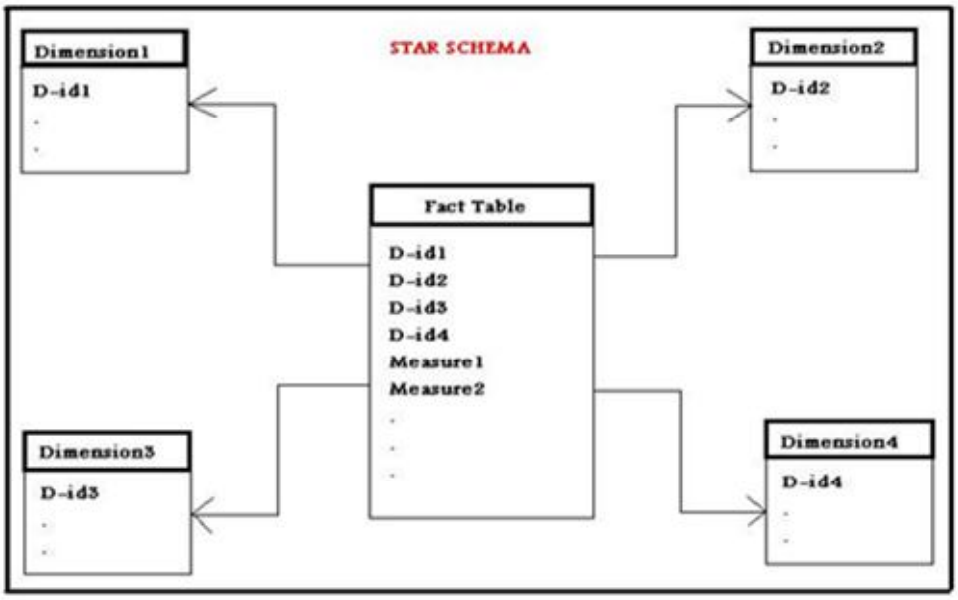


Figure 3-7Star schema(Al-Rammahi 2015)

- *Snowflake Schema*

The snowflake schema is an extension of the star schema. In a star schema, each dimension is normalized and connected to more dimension tables,figure 3-8.

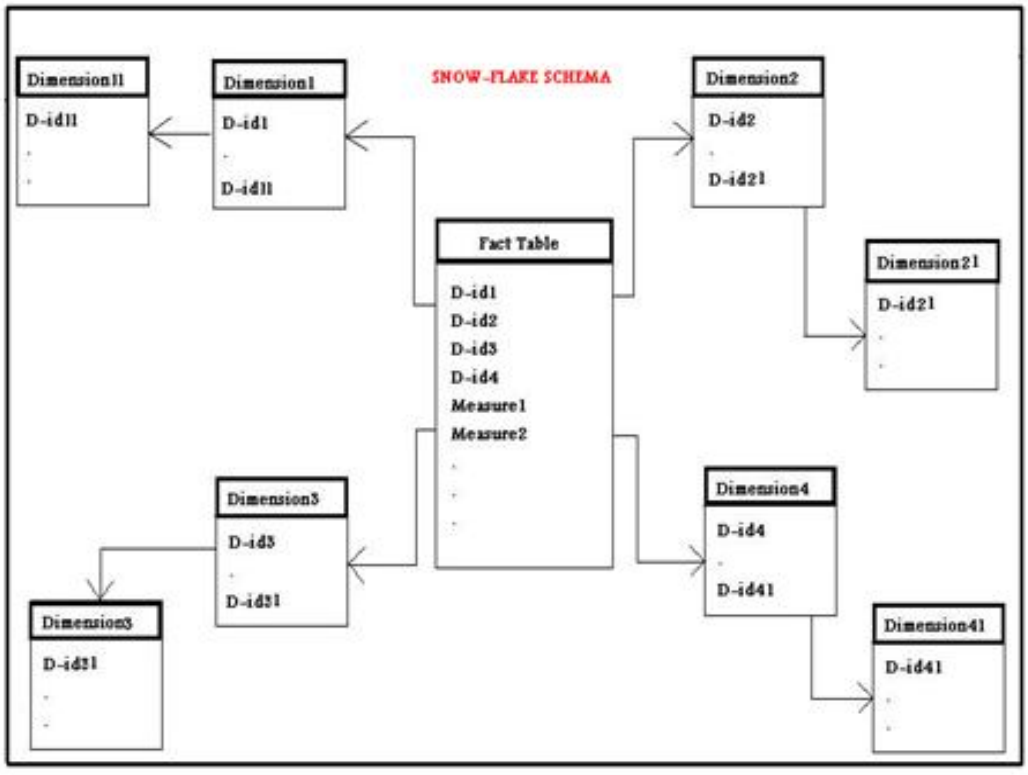


Figure 3-8 Snowflak schema(Al-Rammahi 2015)

3.2.2.2 Business Analytics

Looking at all the data to understand what is happening, what will happen, and how to make the best of it for doing this variety of tools and techniques can be used(Michigan 2017). These tools and techniques fit into two major categories(Turban, Sharda et al. 2014):

- Reports and queries. Business analytics include static and dynamic reporting, all types of queries, discovery of information, multidimensional view, drill down to details, and so on.
- Data mining**:** is a process of searching for unknown relationships or information in large databases or data warehouses, using intelligent tools such as neural computing, predictive analytics techniques, or advanced statistical methods.

  There are several major data mining techniques have been developing and using in data mining projects recently here is description of these techniques(Lau 2003, Olson and Delen 2008):

Classification

Classification is a classic data mining technique based on machine learning. Basically classification is used to classify each item in a set of data into one of predefined set of classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics. In classification, we develop the software that can learn how to classify the data items into groups.

Clustering

Clustering is a data mining technique that makes meaningful or useful cluster of objects which have similar characteristics using automatic technique. The clustering technique defines the classes and puts objects in each class, while in the classification techniques, objects are assigned into predefined classes.

Neural networks

Neural network is a set of connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights

so as to be able to predict the correct class labels of the input tuples. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. These are well suited for continuous valued inputs and outputs.

Association

Association is one of the best known data mining technique. In association, a pattern is discovered based on a relationship between items in the same transaction. That's the reason why association technique is also known as relation technique.

Prediction

The prediction, as it name implied, is one of a data mining techniques that discovers relationship between independent variables and relationship between dependent and independent variables.

Educational Data Mining (EDM)

An international Educational Data Mining community defines EDM as an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in (EDM.org 2011).

The educational data mining process converts raw data coming from educational systems into useful information that could potentially have a great impact on educational research and practice. In addition to data mining techniques Educational data mining apply other techniques like regressions, correlation and visualization(Romero and Ventura 2010).

3.2.2.3 Business Performance Management

Also referred to as corporate performance management (CPM), business performance management (BPM) is an emerging portfolio of applications and methodology that contains evolving BI architecture and tools in its core. BPM extends the monitoring, measuring, and comparing of sales, profit, cost, profitability, and other performance indicators by introducing the concept of management and feedback. It

embraces processes such as planning and forecasting as core tenets of a business strategy. In contrast with the traditional DSS, Executive Information System (EIS), and BI, which support the bottom-up extraction of information from data, BPM provides a top-down enforcement of corporate-wide strategy(Turban 2011).

3.2.2.4 The User Interface

Dashboards provide a comprehensive visual view of corporate performance measures (also known as key performance indicators), trends, and exceptions. They integrate information from multiple business areas. Dashboards present graphs that show actual performance compared to desired metrics(Turban 2011).

3.2.3 Role and benefits of BI

Business Intelligence enables organizations to make well informed business decisions and thus can be the source of competitive advantages.

Improve the timeliness and quality of information. It can eliminate a lot of the guesswork within an organization, enhance communication among departments while coordinating activities, and enable companies to respond quickly to changes in financial conditions, customer preferences, and supply chain operations. BI improves the overall performance of the company using it(Ranjan 2009).

The objective of Business intelligence (BI) is to improve the timeliness and quality of information, and enable managers to be able to better understand the position of their firm as in comparison to competitors(Khan and Quadri 2012).

CHAPTER 4

4 Literature Review

The HE sector today is a dynamic and complex environment with an increasing competitiveness between universities due to the rapid development of information and communication technologies.HEIs that want to stay competitive have realized the need to analyze the available data to thoroughly understand their organization in regards to their students, administrative staff, class and schedule information etc. (Kabakchieva 2015).The adoption of advanced analytical technologies e.g. BI and data mining technologies provide possibilities to extract useful information to enhance decision-making and improve organizational efficiency. In this chapter we try to review studies published about BI utilization and applications at HEIs and data mining application in HEIs.

4.1 BI applications in HEIs

Dell et al presented an approach to implement Business Intelligence systems in one Italian university. The work focuses on how to build an academic data warehouse to support the decisional and analytical activities. The analysis is done for students' data and enrollment data using online analytical processing OLAP and data mining. The system generates many reports as for example the name of the faculty whose students pass the examinations with highest score, the number of graduate students in each college in a specific period of time. The aim of this reports and analysis is to give some indicators about the quality level of the teaching process in the university. However, the tool used to implement the BI is not determined in the paper, also details are missing about data mining techniques used in analyzing the data(Dell'Aquila, Di Tria et al. 2008).

Azman et al developed BI systems for the University Utara Malaysia (UUM) as a case study. The BI system is limited only to business processes in student affairs. In developing BI systems the authors use SAS tools in two stages, known as back room and front room activities. Back room activities refer to the preparation of the data for a data warehouse using the SAS ETL (Extract, Transform, Load) studio 3.3, and front room activities refer to the manipulation and analysis of the data warehouse data for providing information using the SAS Enterprise Guide. Simple reporting is developed to support decision making processes (example: Report about the total number of student

registered for academic session July 2006/2007). However, the work focuses on developing BI for student affairs, and there is no analysis on how the report generated by the system can help in supporting the decision making process itself(Saleh, Ta'a et al. 2008).

Piadade and Santo proposed the implementation of Student Relationship Management (SRM), or Customer Relationship Management (CRM) in higher education. The concept of SRM is adapted from CRM, a CRM is used in a business environment to support and manage the relationship between the organization and their customers. The SRM should support the activities related to students and associated with the teaching-learning process. The SRM architecture contains four components: Data warehouse, data analysis, interaction component and the assessment component. In developing and implementing the BI system they used the SQL Server Business Intelligence Development Studio (MySqlServerBIDS). For building ASP.NET web application they used Visual Studio .NET as tool. The analysis of the students' data is done applying data mining technology. The main purpose of data mining analysis is to identify the students profile in order to decrease the failure rate and the failure risk in

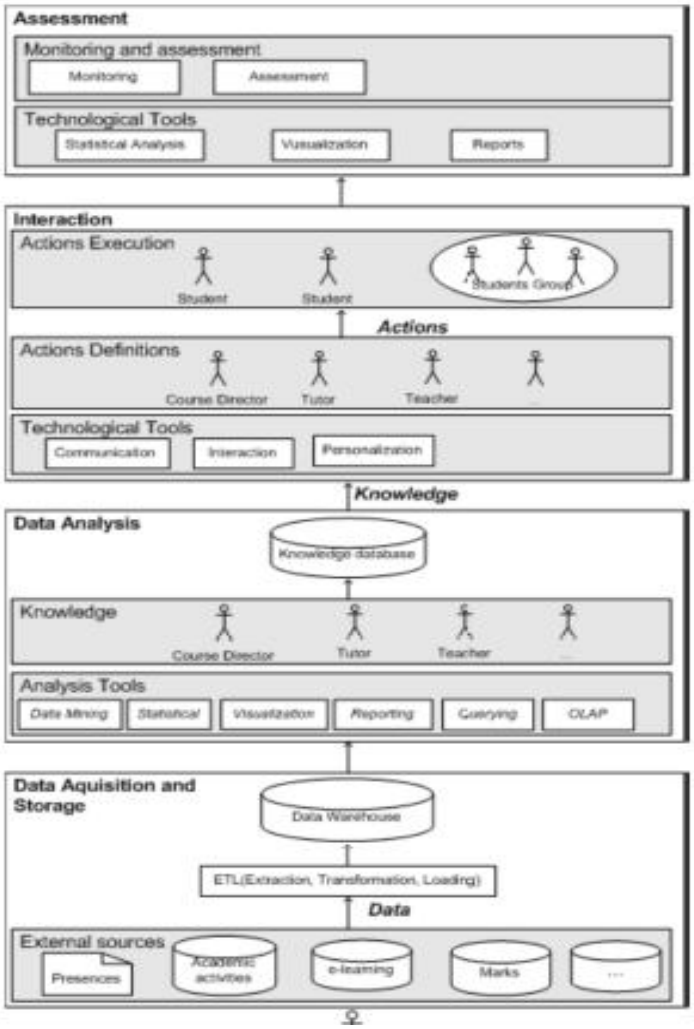


Figure 4-1 SRM architecture (Piedade and Santos 2010)

future(Piedade and Santos 2010).

Azwa Abdul Aziz et al. (Aziz, Rizhan et al. 2012)present an Educational Intelligence (EI) Framework by combining BI technologies with an Educational Data Mining (EDM) algorithm. They take the University Sultan Zainal Abidin as a case study. The data are collected from students' examinations result. The analysis is done using association rules as part of the predictive technique. The authors applied the Talend Open Studio (TOS) as tool to ETL and IBM Cognos BI tool. The analysis of the

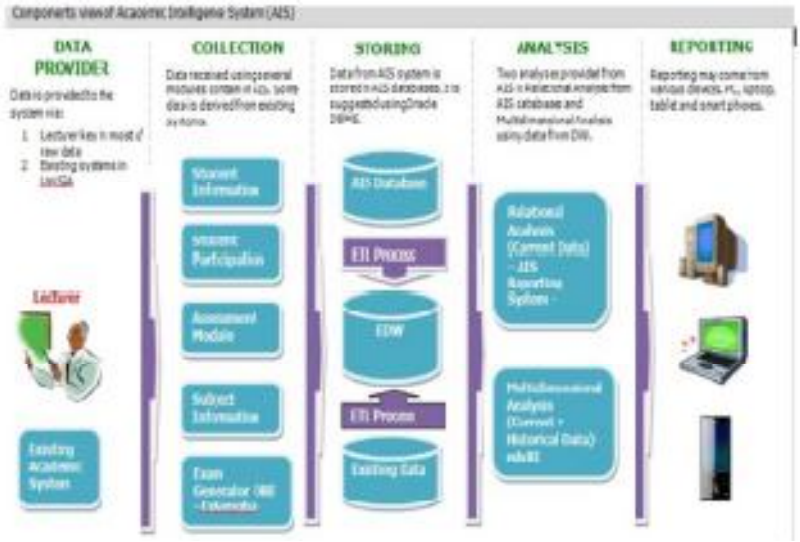


Figure 4-2 EI framework (Aziz,Rizhan et al. 2012)

information serves to derive the best teaching mechanisms for an individual student.

Another study, provided by Mohiuddin Ali et al. presents a BI system for the analysis of the students' results using EDM. The analysis is based on clustering techniques and is based on the Apriori algorithm. The data are collected from number of institutions. The authors observed that there is some relation between failure in English and another courses, and when a student has failed in English and mathematics he/she also failed in programming subjects. However, there is no description of the data set and also the tools are not determined(Khan, Gharibi et al. 2014).

Kabakchievain in her studypresents an example of BI implementation for student data analysis. The data collected are by one lecturer about exercise classes within a university course taught at the university of National and World Economy in Sofia, Bulgaria. The data set contains 575 records in three academic years. QlikView is the

software used for developing the BI application. The system generates information used to conduct the student performance analysis. The work shows how the BI solution can advantageously provide information to support the decision making process for educational purposes. However the data used is very limited. If the data analyzed will increased, the analysis will give more accurate results(Kabakchieva 2015).

Zhang et al. proposed a system based on data mining where different data sources from the Thames Valley University systems are used. Data sources covers student enrolment, students result, course/module data, learning skills, and student activities. This data sources are integrated and transformed into data warehouse. The data from data warehouse were used by the data mining engine, three different algorithms used for students' dropout prediction: Naïve Bayes, Support vector machine and Decision tree. Text mining and natural language processing technology are applied to user friendly generate results that can be easily understood. To build the system, Oracle 11g is used as a platform integrated with an Oracle data warehouse builder and Oracle data miner(Zhang, Oussena et al. 2010).

Zilli (Zilli 2014)presents according to Gartner a definition for the concepts of a self-service BI system as end users designing and deploying their own reports and analyzes within an approved and supported architecture and tools portfolio. He also discuss the methodology for proposed BI system, and different Key performance indicator (KPI) used in higher education. The main KPI's discuss in study are: students' enrollment, exam performance rate, course work, under graduate retention rate, graduation rate and academic work load. The author present a dimensional model of proposed BI solution. He used Excel PowerPivot modeling tool in designing a dimensional model.

Falakmaisr et al. presents the use of BI and online analytical processing (OLAP) tools in an e-learning environment. The e-learning department of the Iran University of Science and Technology (IUST) served as a case study. This department used Moodle as a learning management system, the features of using Moodle that it's an open source system and also another important feature of Moodle that it stores logs in relational database rather than text file. The authors build a data warehouse and apply OLAP as analytical method. The study examined activity logs of 1300 students in 100 courses over 16 weeks. The result of the analysis shows that students spend little time with

course courseware and prefer to use collaborative activities, such as a virtual class room and forum instead of just learning materials(Falakmasir, Moaven et al. 2010).

Abdul Aziz et al. proposed educational data warehouse architecture. The authors integrate proprietary and open source BI tools. The educational data warehouse framework generates set of reports. An example of reports is the students' result by subject and geographic information(Aziz, Jusoh et al. 2014).

Julaily Aida et al discuss in their paper designing a conceptual framework for business intelligence analysis in academic environment, Figure 4-3. The framework divided into three phases, input phase, process phase and output phase. They used dummy data for analysis; they assumed that the data collected from two-sub system assessment and allocation system. They used this data to determine and analyze the correlation between students result and accommodation. In the output phase they used Business Intelligence and Reporting Tool (BIRT) to represent the reports. The graph generated by BIRT tool shows that the most in-campus students achieved better results compared to out-campus students(Julaily Aida, Raja Hasyifah et al. 2012).

Pidade and Santos presents application case depends on use student relationship management concept. The data gathered from a graduate course of a Portuguese higher education institution. The data is for 70 students in the first year in e-learning platform, this data loaded from different database into data warehouse. OLAP techniques used to analyze the student's results verifying both the teaching – learning experiences and the assessment methods. Also data mining algorithms were used to identify the students' profile considering the assessment result. Decision tree was used as classification method(Piedade and Santos 2009). However the data set used in this study is very small

Wan maseri et al presents system to generate a novel digital dashboard for HEI, the system process start from capturing data from university database, followed by data clustering process to cluster the students' performance data, then used visualization to

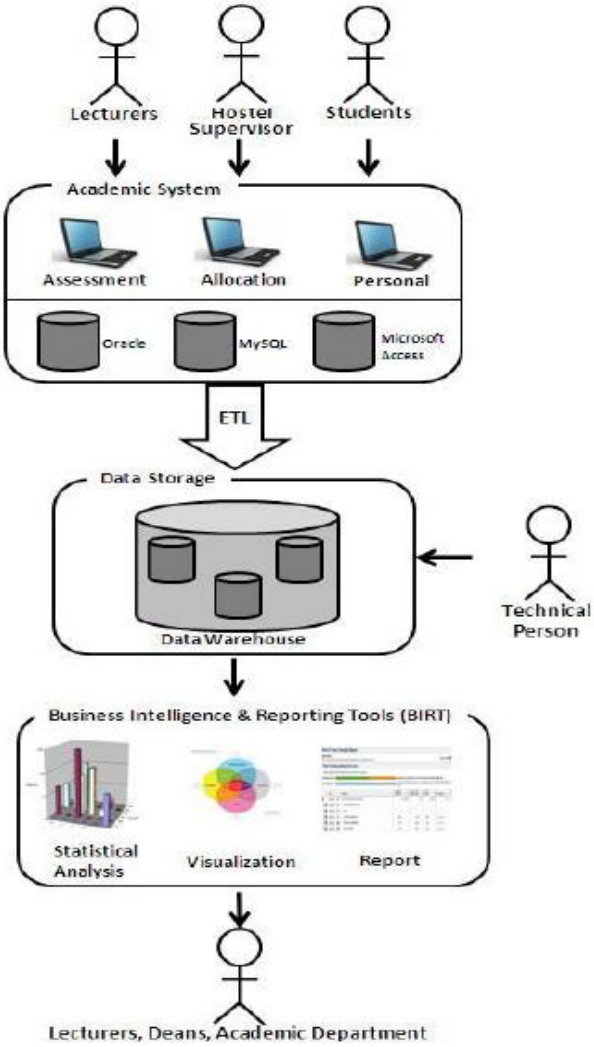


Figure 4-3 BI conceptual framework(Julaily Aida, Raja Hasyifah et al. 2012).

visualize the generated clusters. And for predicting the future performance they used fuzzy pattern matching algorithm. The system finally used the concept of dashboard to present the result. The first segment of dashboard shows the current status of students' performance followed by forecast performance and previous performance. The dashboard help manager to have general view about students' performance and this will support the decision making process(Mohd, Embong et al. 2008).

4.2 Data mining application in HEIs

Yadav and al. propose a data mining methodology to select students for an enrollment in a particular course. The authors apply classification techniques to evaluate the previous student's performance, the classification method used is decision tree. The data of 430 students collected, the data about demographic data, past performance data, address and contact number. The model constructed depend on ID3 decision tree. The accuracy of the model is 60.46% from 430 instance 260 classified correctly. The data mining tool used is weka (Yadav and Pal 2012).

Bahardwaj et al apply EDM to analyze the students' performance based on classification techniques. The data set used in this study was obtained from the Purvanchal University, from 2007 to 2010. The authors collect data about the attendance, class tests, seminars and assignments from a students' earlier database, the data collected for 50 students, to predict the performance at the end of the semester. The classification method used is decision tree(Baradwaj and Pal 2012).

Borkar and Rajeswari used association rules to evaluate the students' performance. The data set covers 60 students from a master course in computer application, Pune University. Weka has been used as a data mining tool. The study reveals that the student's performance depends on the test, assignment, and attendance and graduation percentage. The university can improve the result by giving additional guidance to the students who poorly performed for example in unit tests, assignments, and attendance and graduation percentage(Borkar and Rajeswari 2013).

Durairaj and Vijitha exploitEDM for predicting the students' performance based on clustering and classification as data mining techniques. They use a sample of 38 students' record for the data analysis, and used weka as a data mining tool(Durairaj and Vijitha 2014).

Harilatha and Sudhakaryadv present data mining techniques to predict failure and dropout of students. They used classification methods for classifying real data of students. Java programming language is used for the attributes selection, pre-processing and classification. The result of this study is that the dropout by male and female differentiation is based on given input data. However, the attributes used in the classification are not determined and the data available is imbalanced (between male and female)(Harilatha and Sudhakaryadav 2014).

Pal used data mining techniques to reduce dropout rates of engineering students, he used classification methods. The data set consist of 1650 record, the data is collected from the Institute of Engineering and Technology, Purvanchal University. The results show that the data mining technique is able to establish an effective predictive model from the existing student dropout(Pal 2012).

Kumar and Chadha presents the use of data mining techniques in analyzing the huge volume of data in higher education, they presents number of applications of data mining: the first one is organizing the syllabus, this can done by identify relate subjects in syllabus, the suggest to use association rule mining in identify possibly related two subject. The second application is predicting the registration of students in an educational program, they present that if the number of students is predetermine the resource of organization can allocated effectively, they suggest that using the prediction techniques can help in accurate estimation of how many male or female will register. The third application is predicting student performance, the authors discuss the important of evaluating the students' performance and presents that the data mining techniques can be used in develop a result evaluation system, the system can help teachers and students to know the weak points of the learning process and try to face the problem. The appropriate data mining techniques to use in this application is classification. The fourth application is detecting cheating in online examination, data mining techniques can propose model which can help organizations to detect and prevent cheats in online assessments. The last application is identifying abnormal/erroneous values; data mining can be used to detect the abnormal values in the result sheet of students (Kumar and Chadha 2011).

Singh et al in their paper presents use of data mining to extract information from available data in educational institution. The students data used in their work is from database management system coerce held at the united college of engineering and research Naini, Allahabad in fourth semester of 2011/2012. And also they used questionnaire to collect the real data that describing the relationships between learning behavior of students and their academic performance. In their study they used association rules, classification, and clustering as data mining techniques. However the discussion of result not presented (Singh, Tiwari et al. 2013).

Dorina Kabakchieva present in her paper the initial result from a data mining research project, the project implemented following the CRISP-DM (Cross Industry Standard Process for Data Mining) model which is cycle approach consist of six main phases: business understanding, data understanding, data preparation, modeling, evaluation and deployment. The software used is Weka, the data used consist of about 10330 students' record described by 20 parameters. In this study different classification algorithm: decision tree algorithm, Bayesian classifier, Nearest neighbor algorithm and rule learners(Kabakchieva 2013).

Mohd maqsodd ali present research paper about data mining and its role in education sector, he discuss different suggested data mining application in education sector, benfit from using data mining, data mining techniques and author also discuss data mining process from different authors perpective. The applications presented by the author are: students' enrolment prediction, predicting students' profiling, curriculum development, students' complaints, course completion, students' targeting, library facility, students' course selection, teachers' teaching performance, students' performance, students' dropouts and students' relationship management(Ali 2013).

E-khadr and Elsddawy proposed system called Higher Education Data mining system, this consist of several steps: data gathering, preparing data, data preprocessing, using data mining techniques, post processing and get result and visualization. The data mining techniques used in this framework are clustering (used K-means and Enhanced K-means method), classification and association rules. The data collected from Moodle system. Different software used: Microsoft Excel used for analysis and filtering data, Matlab for data preprocessing and data classification and finally weka(Khedr and El Seddawy 2015).

R.B Bhise et al presents a study of using K-means as clustering method to analyze students' data in order to show how different factor affect a student learning behavior and performance. The data gathered from department of bachelor of computer application, Ale University. The software used for developing the system was visual studio 2008. However the details about data set not presented(Bhise, Thorat et al. 2013).

Table 4-1 summary of BI application

Paper	Operational system	Data used for analysis	Results	Software used	Year & author
Academic business intelligence system development using SAS tools	Graduate academic IS Students affair IS Personnel IS	Total number of students	BI system generate reports about students registered by programs	SAS	(Saleh, Ta'a et al. 2008)
Business Intelligence Systems for Analyzing University Students Data	Personal records of a lecturer teaching exercise classes for three different years	Gender, year , number of absence, access score, excel score, final score	There is a relation between student performance and absence Female students performance better than male and this valid for three years	QlikView	(Kabakchieva 2015)
Conceptual framework of business intelligence analysis in academic environment using BIRT	Assessment system Allocation system Personal system	GPA for one semester, living in or out campus	There is a correlation between an accommodation and students results.	BIRT	(Julaily Aida, Raja Hasyifah et al. 2012).

Self-service BI system to manage HEIs and enhance the efficiency of the study process.	Data about students, staff, and faculty and also data related with KPIs: students Enrollment, Exam Performance Rate, Retention Rate, and Faculty's Academic Workload.		Design dimensional model for Study process
Business Intelligence Applications for University Decision Makers	six source databases: * Secretary Services for Students (ESSE3) DB that supports didactic curricula, admin processes & services to students. * NOt managed (NOGE)) a legacy DB that contains residual historical data about students enrolled before introducing ESSE3. * Athenaeum Integrated Accounting (CIA) financial		Building data mart for departmental areas: Didactics, Finance, Research & HR.

	mgmt system that views the university as a business. * Careers & Wages of Athenaeum (CSA) takes care of the legal and economic management of the university personnel. * SAPERI is the DB of the scientific research competence of the University. It includes publications & patents of researchers. * SINBAD for the management of the athenaeum research projects.		

A framework for educational data warehouse (EDW) architecture usingbusiness intelligence (BI) technologies	A set of data from students' examination results in relational database		* Design multi-dimensional model * Set of reports about students' result by subject, geographic information, time dimension and personalized students' achievement.
Promoting students sources with a business intelligence system	* Adopted learning process *Assessment method *Development activities *Student and their interaction in developed activities.		* Defined the SRM concept. * data mining analysis for data in order to decrease the failure rate and the failure risk in future
Business intelligence in e-learning:(case study on the Iran university of science and technology dataset)	Moodle logs data	1300 students in 100 courses over 16 weeks	*Using OLAP result of the analysis shows students : * spend little time with course courseware. * prefer to use collaborative activities

	*student enrolment * students result *course/module data * learning skills *student activities		
Data mining techniques for business intelligence in educational system: A case mining	*Students data *Courses information *Academic data		* Student who failed in English course have also failed in many other courses *Student who failed in English and mathematics have also failed in programming subjects.

CHAPTER 5

5 Methodology

This chapter describes the research methods and analytics technique used in order to achieve the objective of the research.It's organized into three phases: in the first phase content analysis, analytical research procedures, and exploratory interviews were used while performing an extensive review of literature to identify the factors that affect the learning process (chapter two), stakeholder, and data sources required for building the BI solution. The second phase questionnaire designed to gauge the effect of the previously identified factors on the process. The third phase explores the development of a BI model that takes advantage of the identified learning process influences and the data collected.

5.1 Needs analysis and requirements gathering

5.1.1 Participants

The participants in this study are university students. The students are from college of computer science and information technology (CSIT) in Sudan University of science and technology (SUST), from four departments: computer sciences (CS), computer and information system (IS), computer systems and networks (NW) and software engineering (SE). The number of students who responded to questionnaire are 347, the general characteristic describe in table 5-1 below. 153 are male (44%) and 194 are female (56%), figure 5-1.

Table5-1 General characteristic

Department	Number of students	Male	Female
CS	83	36	47
IS	72	35	37
NW	79	28	51
SE	113	54	59
Total Number	347	153	194

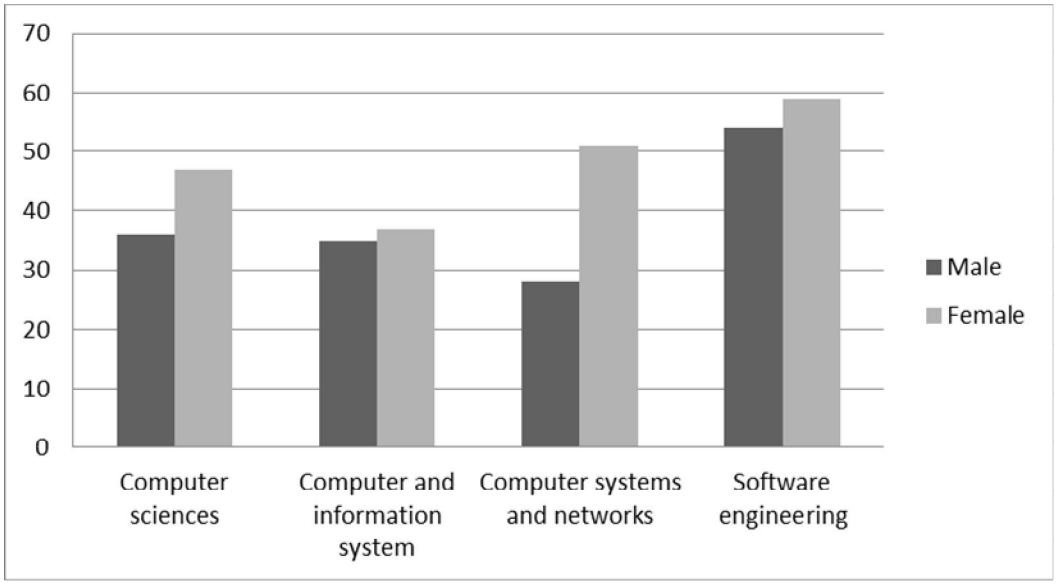


Figure 5-1 Gender

5.1.2 Questionnaire Design

The questionnaire passed through a number of stages until it reaches its final form, the stages are:

- Select the Factors
- Prepare first version of questionnaire to collect the data.
- Submit the questionnaire to the supervisor in order to determine the suitability of it to data collection.
- After discussion and many change we choose final form
- Translate the questionnaire to Arabic, because Arabic is mother language of students and students will understand the questions
- Choose number of students (testing case) to discuss how they understand the questions, is questions are clear or there is problems.
- After all this stages the questionnaire is ready and distributed to students.

The questionnaire contains two sections:

- Section one:

  Is about personal data to illustrate the characteristics of the sample of the study, include the college, department, academic year and gender.

- Section two:

  The questionnaire consists of 71 sentences (Appendix A) about the factors that affecting the learning process and is divided to sub section as follow:

  - Environment

    Contain 18 statements which discuss the university environment inside and outside classroom.

  - Teaching

    Contain 21 statements which discuss the timetabling, teaching methods, teaching support, communication, support for collaborative work and examination and assessment methods.

  - Academic advising

    Contain 3 statements which discuss the availability of adviser and if they support and help in determining goals.

  - Student

    Contain 11 statements which discuss family support, personality and activities.

  - HEI

    Contain 10 statements which discuss curriculum design, additional education opportunity, support for students and students' affairs.

  - ICT support

    Contain 8 statements which discuss technical support, availability of network and internet and availability of Learning Management System.

The five Likert scale was used to measure student responses to the questionnaire sections. The Likert scale is one of the best types of measurements when we want to collect specific data about factors contributing to a particular feeling.

Most studies encourage the use of five alternatives that contain a neutral point in the middle because when the number of alternatives is less than five, the respondent is forced to agree or disagree, so the results do not express the point of view of the person strongly. When the number of alternatives is more than five, it is difficult to interpret the results(Leung 2011).

5.1.3 Technique for analysis

Clustering is unsupervised learning data mining techniques; it is a process of partitioning a set of data (or objects) into a set of meaningfulgroup, called clusters. Clustering algorithms can be used to find natural groupings when there are many cases and no obvious groupings. Clustering relevant data in different clusters can be easy to analyze, in this study we need know about the factors that negatively affect the learning process from students view to be improved so partitioning students in groups according to their characteristic and how they asses this factors make it easy to analyze the data and understandthe problems.

K-means algorithm is most widely used algorithm for clustering data in practice, its hard clustering algorithm; each object is assigned to one cluster. The k-means algorithm is simple and easily understandable(Berkhin 2006).

5.1.4 The algorithm

Algorithm: *k*-means. The *k*-means algorithm for partitioning, where each cluster's center

is represented by the mean value of the objects in the cluster.

Input:

k: the number of clusters,

D: a data set containing *n* objects.

Output: A set of *k* clusters.

Method:

(1) arbitrarily choose *k* objects from *D* as the initial cluster centers;

(2) **repeat**

(3) (re)assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster;

(4) update the cluster means, that is, calculate the mean value of the objects for each cluster;

(5) **until** no change;(Han, Pei et al. 2011)

The similarity between object is determine here according to distance function, in our study we us the most common one Euclidean function.

The k-means algorithm is appropriate for mining educational data because it's one of fastest partitioned method for clustering large data.

The WEKA data mining tool used for mining the data, WEKA is refer to**W**aikato **E**nvironment for **K**nowledge **A**nalysis, data mining/machine learning tool developed by Department of Computer Science, University of Waikato, New Zealand(Bouckaert, Frank et al. 2010). WEKA is an open source tool,it supports many different standard data mining tasks such as data preprocessing, classification, clustering, regression, visualization and attribute selection.WEKA is one of ease of use tool and also many built-in features that require no programming(Rangra and Bansal 2014).

5.1.5 Steps for analysis

1. Coding questionnaire data in excel file.
2. Convert file to CSV format, a comma-separated values (CSV) file is a delimited text file that uses a comma to separate values. A CSV file stores tabular data (numbers and text) in plain text. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format. We need this step because WEKA support CSV file.
3. Open file in WEKA environment and choose K-mean clustering algorithm and change the variables according to need. Choosing the number of clusters (K) in k-means depend on user, in this study after testing different number of clusters, three clusters (k=3) give clear and good result. The students divided in three different groups, according to pattern of answers, the name of groups are cluster0, cluster1 and cluster3.

Another techniques for analysis is using traditional method (calculating Mean) for all variables (Factors). To perform this, data from the questionnaires was decoded and captured into IBM's SPSS for Windows.

5.2 Data analysis

5.2.1 Environment:

Table5-2 Environment dimension

		K-means			Statistical analysis	
Factor		**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
1	HEI infrastructure	2.3	3.3	2.8	2.8	Neutral
2	Transportation	1.6	2.5	1.8	2.5	Disagree
3	Classroom environment	2.2	4.0	3.1	3.3	Neutral
4	Study environment	2.0	3.5	2.9	2.9	Neutral
5	HEI services	1.8	2.9	2.4	2.5	Disagree

Table 5-2 show the results of the five Environment factors and the Mean for the three clusters, i.e., cluster0, cluster1 and cluster2. Each cluster represents a group of students and the Mean is the output from the K-means algorithm. The Mean here represents the weighted average. Since we are using Likert quintet, the range is 4, and the number of options equals 5, then the weighted average becomes 0.8. Therefore, each option Mean falls in the following range: 1-1.79 “strongly disagree”, 1.80-2.59 “disagree”, 2.60-3.39 “neutral”, 3.40-4.19 “agree”, and 4.20-5 “strongly agree”. It is clear from Table 5-2 that “Transportation” is a problem that faces the majority of students. All students have to commute to and from the SUST campus and the lack of transportation affects their

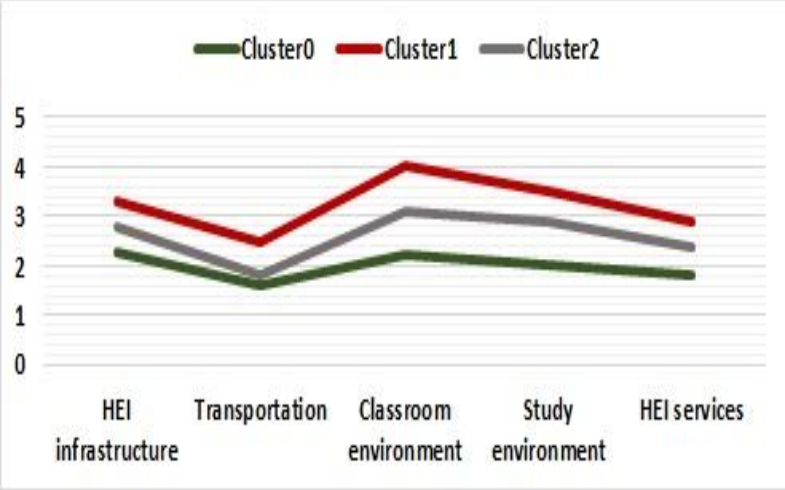


Figure 5-2 . Environment Dimension

morning lectures attendance. The results of “HEI services” too show that most students are not satisfied with services such as healthcare and cafeteria services.

From statistical analysis the students dissatisfied with transportation and HE services the mean value is 2.5 for the two factor.

Figure 5-2 is a line graph of the Environment dimension results, where the y-axis represents the Mean. From the figure we can clearly see that students grouped in Cluster0 are not satisfied with the environment factor in general. Cluster1 students are mostly neutral but they are satisfied with the classroom environment. Cluster2 students are dissatisfied with both transportation and HEI services while they are neutral with other factors.

5.2.2 Student:

Table5-3 Student dimension

	K-means			Statistical analysis	
Factor	**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
Family support	3.7	4.4	4.1	3.9	Agree
Personality	2.6	3.8	3.4	3.2	Neutral
Extracurricular	1.7	3.4	1.9	2.5	Disagree

Table 5-3 presents the results of the Student dimension. “Family support” explores the student’s relation with his/her parents, the ability to communicate with

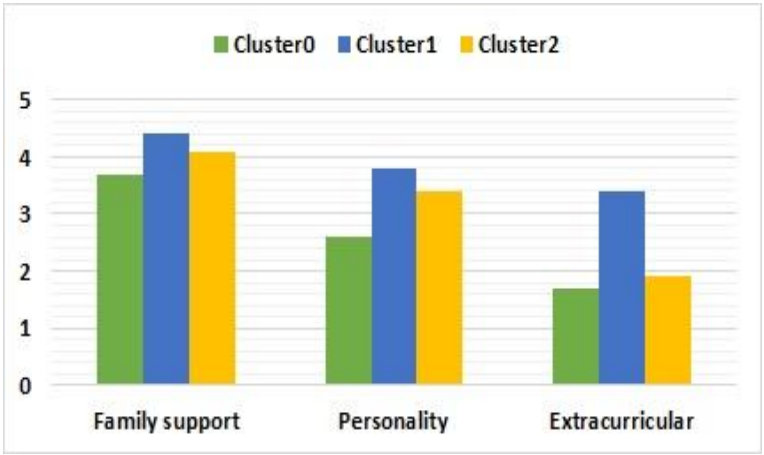


Figure 5-3Student Dimension

them, and also if the student can pay tuition fees. From Table 5-3 results, the general pattern of answers is “agree”. This means that most students have good financial and moral support from their families.

Figure 5-3 is a bar-chart representation of the Student dimension and the y-axis represent the Mean. The second factor, i.e. personality, exposes the students’

desire to study, communicate with colleagues, and communicate with people responsible of his/her affaires. From the results in Figure 5-3, we can see that some of students have problems with this factor. The third factor deals with extracurricular activities such as political, sports, cultural, etc. It addresses the availability of time and infrastructure to practice such activities on the SUST campus. The results in the table and graph clearly demonstrate that SUST needs to put more emphasis on supporting these activities as they positively affect students and the learning.

5.2.3 Teaching

Table5-4 Teaching Dimension

	K-means			**Statistical analysis**	
Factor	**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
Lectures Timetable	1.4	2.8	1.9	2.5	Disagree
Language barrier	2.3	3.4	2.8	2.9	Neutral
Teaching method	2.3	3.4	2.8	2.6	Neutral
Teaching support	2.8	4.1	3.6	3.4	Agree
Communication	2	3.5	2.8	2.6	Neutral
Collaborative work	2	3.8	3.2	3.1	Neutral

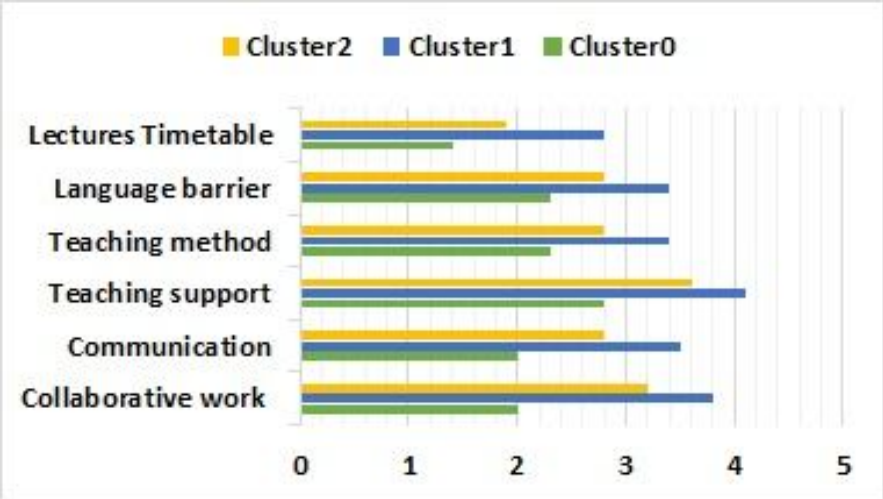


Figure 5-4 Teaching Dimension

Result of the analysis of the Teaching dimension and its factors is presented in Table 5-4 and Figure 5-4. Generally, most responding students ranked all factors low except

the students in Cluster1, they ranked “Teaching support” and “Collaborative work” high. These results reflect the strong dissatisfaction of students with the teaching dimension, and have negative impact on the learning process. Consequently, there is an urgent need for improvements as the teaching factors play vital role in learning process and have direct relation with students’ satisfaction.

5.2.4 Academic Advice:

Table5-5 Academic advice

	K-means			**Statistical analysis**	
Factor	**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
Office hours	1.2	3.7	2.2	2.4	Disagree
Communication	3.5	3.4	1.7	2.5	Disagree
Help to determine goals	3.2	3.6	1.5	2.4	Disagree

With regard to this dimension, students were asked about the office hours for meeting their advisor, the communication with the advisor, and if they get adequate assistance from her/him to achieve their goals. From the result presented in Table 5-5, there are varying views between dissatisfied and satisfied, and this is due to the evaluation of different advisors. This indicates that there is a difference in performance between academic supervisors.

From analysis the students not satisfy with this factor and there is a need to improve the academic advice process.

5.2.5 HEI:

Table5-6 HEI Dimension

	K-means			**Statistical analysis**	
Factor	**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
Curriculum design	2.9	2.4	2.0	2.4	Disagree
Additional education opportunities	3.9	1.9	1.9	2.4	Disagree
Financial support	3.4	3.0	1.9	2.7	Neutral
Students’ affairs	3.7	3.3	1.9	2.9	Neutral

HEI dimension has four factors as shown in Table 5-6. The “Curriculum design” factor addresses the process of course design and the sequencing of courses. From the results in Table 5-6, the majority of students are not satisfied with this factor since the Mean value is less than 3 in all clusters, meaning that it negatively affects the learning process. The second factor in the table explores if the HEI provides additional education opportunities, such as professional or summer training, internship support, international research projects, etc. This factor too has a negative effect on learning process because students in Cluster1 and Cluster2 are not satisfied. The third factor in this dimension, deals with the financial support for students. Here responses vary, as some students see that SUST provides enough support while others disagree. The last factor addresses students’ affairs, i.e., do students know where to go if they have a problem or issue, and to whom do they talk. Again, in this factor there is no agreement. From the discussed results we conclude that there are some factors in the HEI dimension that need improvement.

5.2.6 ICT:

Table5-7 ICT Dimension

	K-means			**Statistical analysis**	
Factor	**Cluster0**	**Cluster1**	**Cluster2**	**Mean**	**Result**
Technical support	4.0	4	2.1	2.9	Neutral
Network and internet Availability	3.6	2.6	2.2	2.5	Disagree
Availability of LMS	3.6	1.8	2.1	2.3	Disagree

From Table 5-7, most students are satisfied with the first factor, i.e., availability and accessibility of computers and technical support. At the same time, most are not satisfied with the second factor, i.e., internet, network, and Wi-Fi availability, the Mean value is less than 4 in all clusters. This implies that there is a need to provide better internet connectivity with high speed and Wi-Fi connection in order to improve the learning process. The last factor is about the availability of a Learning Management System (LMS), and the majority of students don’t know about LMS and have not used it. ICT is important factor, and there is relation between ICT and other factors. Supporting and encourage the using of ICT will lead to improve the learning process.

5.3 Summary

Analysis of the survey results has shown that there are various factors that adversely affect the learning process at SUST. For example, transportation, HEI services, including cafeteria and healthcare services, lectures timetable, academic advice, curriculum sequences and ICT support. These factors require further in-depth study, focus and analysis by SUST administrators to explore ways to improveperformance and thus enhance the learning process.

CHAPTER 6

6 Proposed BI solution

There is a need to build BI framework to analyze data about the learning process, how it goes to have general picture about the factors to determine if there is any negative point to make an improvement. The mission of this research is to build conceptual framework to store reliable information for analysis and generate reports for support decision making process.

6.1 Academic BI Framework

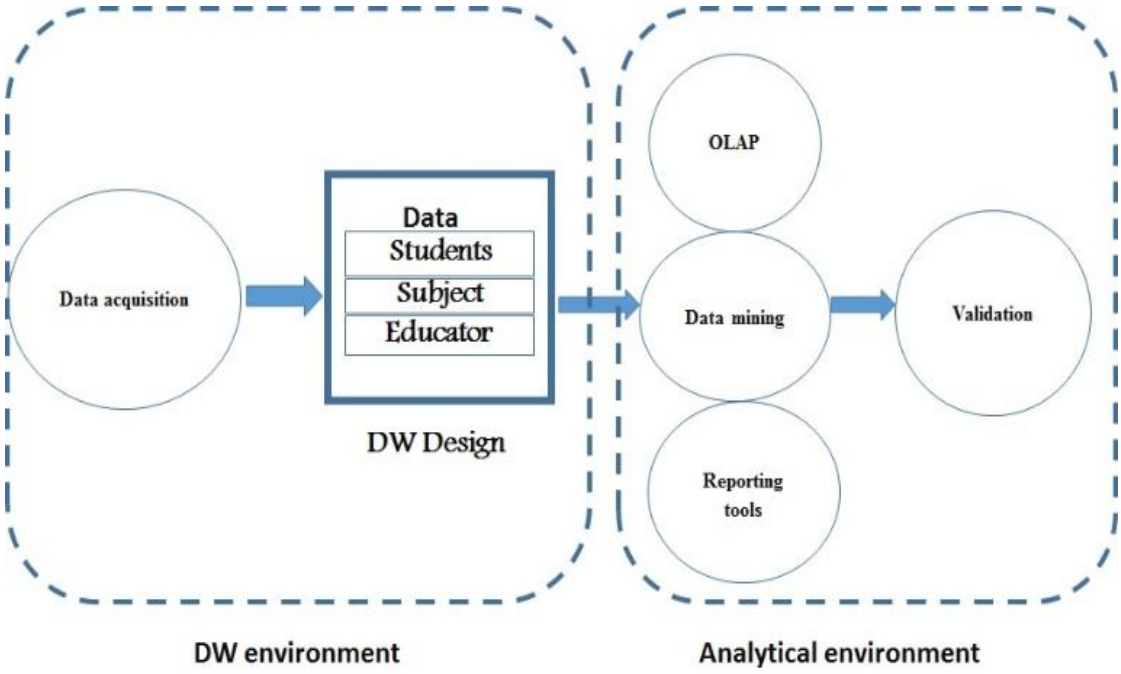


Figure 6-1 Proposed BI solution

Figure 6-1 illustrates the main components of the proposed BI framework that is comprised of two environments, DW and analytical. The DW environment contains the operational systems that supply the core data, the Extract, Transform and Load (ETL) process, and the DW creation. The analytical environment is where the user creates reports and visual aids, queries the system for specific requests, and makes decisions based on information from the DW.

6.1.1 Data acquisition

Based on the factors influencing the learning process outlined earlier in the chapter 2, the data for the proposed BI model comes from eight operational systems at

SUST, as shown in Figure 6-2. Interviews with key personnel at each data source department were conducted to clarify understanding of the data.

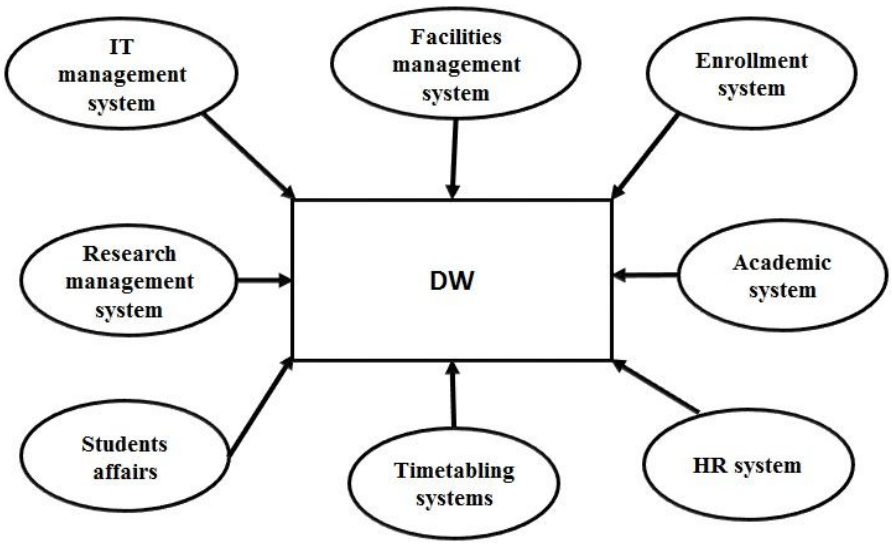


Figure 6-2 Operational systems

First, is the facilities management system that encompasses data about all facilities including class and lecture rooms, laboratories, libraries, offices, cafeterias, etc. Second is the enrollment system that contains students' demographic information and registration data. Third is the academic system from which all data related to courses and students' performance is acquired. Fourth system is the human resource system that covers educators' data. Fifth is the timetable system which comprises the schedules of lectures, seminars, etc. Sixth system is students' affairs which has all the financial support data. Seventh is the research system that contains data about the many research projects at SUST. Finally, the last system is the ICT support system that includes data about computers, internet connectivity, applications including the Learning Management System (LMS), and all technology support related data. The proposed BI framework integrates this carefully chosen datafrom all these sources into a central repository to be used by the appropriate stakeholders.

Table6-1 Data sources

Data about	Factor	Operational systems
Campus, class and lecture rooms	Environment	Facilities management
Academic year, semester, student, college, department	Teaching Advising Student	Enrollment
Course, educator, course assignment, assessment method, references, mark, advisor, student GPA		Academic
Educator		HR
Timetable		Timetable
Financial support	HEI	Students
Research		Research management
Computers, internet speed, WI-FI availability, LMS	ICT	IT management

Table 6-1 displays a mapping of the data sources details to their corresponding learning process factors.

6.1.2 Data warehouse design

6.1.2.1 Logical and Entities relational models

In the beginning the general logical model was prepared which laid out structure of entities, attributes and relationships between entities. SQL Developer Data Modeler (Oracle) used to design this model, Figure 6-3 present the logical model

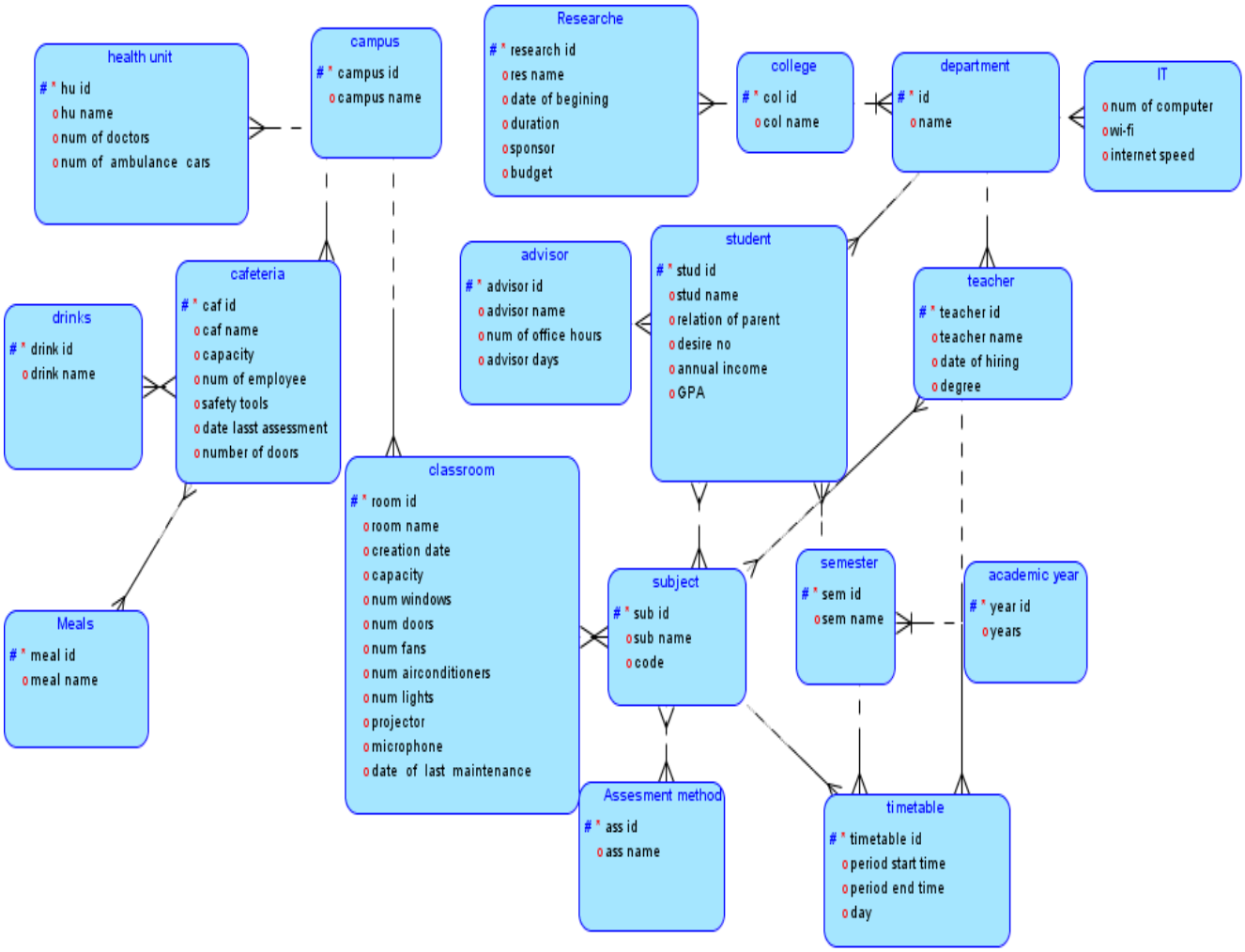


Figure 6-3 Logical model

After the logical data model was completed, then create the ER diagram or physical model using SQL Developer Data Modeler as shown in Figure 6-4 below.

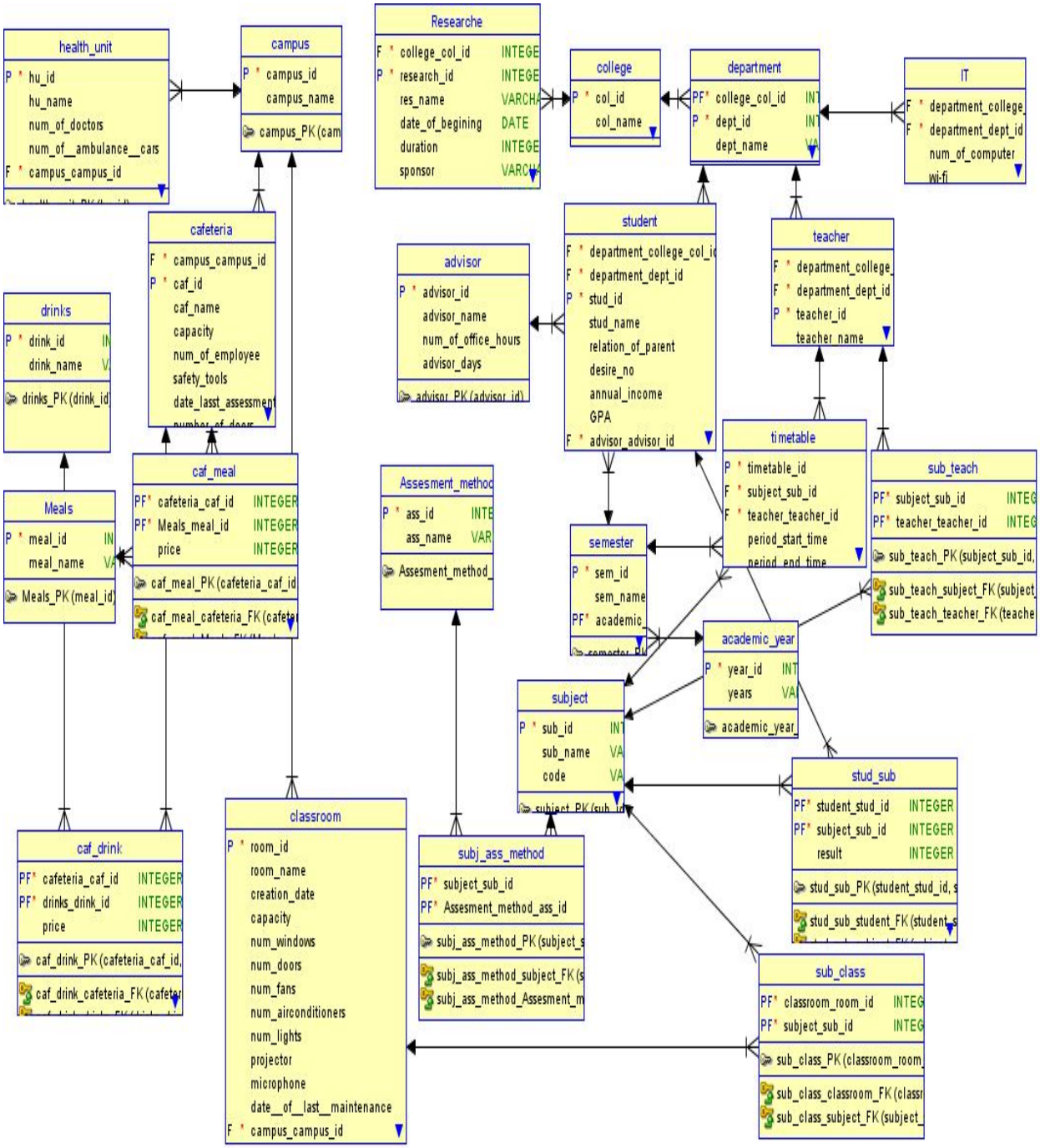


Figure 6-4 ER model

6.1.3 Multidimensional Model

After general specification of each entities, their attributes and relationships the essential step is the design of data Marts which based on designing the dimensional model that related to specific predefined business process.Kimball's data marts are more appropriate for this model (Kimball and Ross 2011), as their staggered approach satisfies the business requirements of the framework and is yet flexible and scalable, i.e., it grows with the availability of additional data.In this study we focused on learning process, so the scope is to build data mart that contain the data of business process related to this issue.

To meet the objective of the research, this case study implementation targets four learning process subject areas, namely, Course-Progress, Class-Occupancy, Financial Supportand Students-Registration. These processes cover the Environment, Teaching, Student and the HEI factors. Exploratory interviews with key stakeholders from the earlier identified SUST departments helped in highlighting and prioritizing these recognized factors. As well, KPIs, possible queries and information dashboard were created based on their feedback so as to study the possible correlation between the factors, their significance and thus validate the framework. Each of the four target subject areas is represented with a data mart that is composed of one central table with a multipart key called the Fact Table and a set of smaller related tables called Dimension Tables. The data marts are:

6.1.3.1 Course-Progress

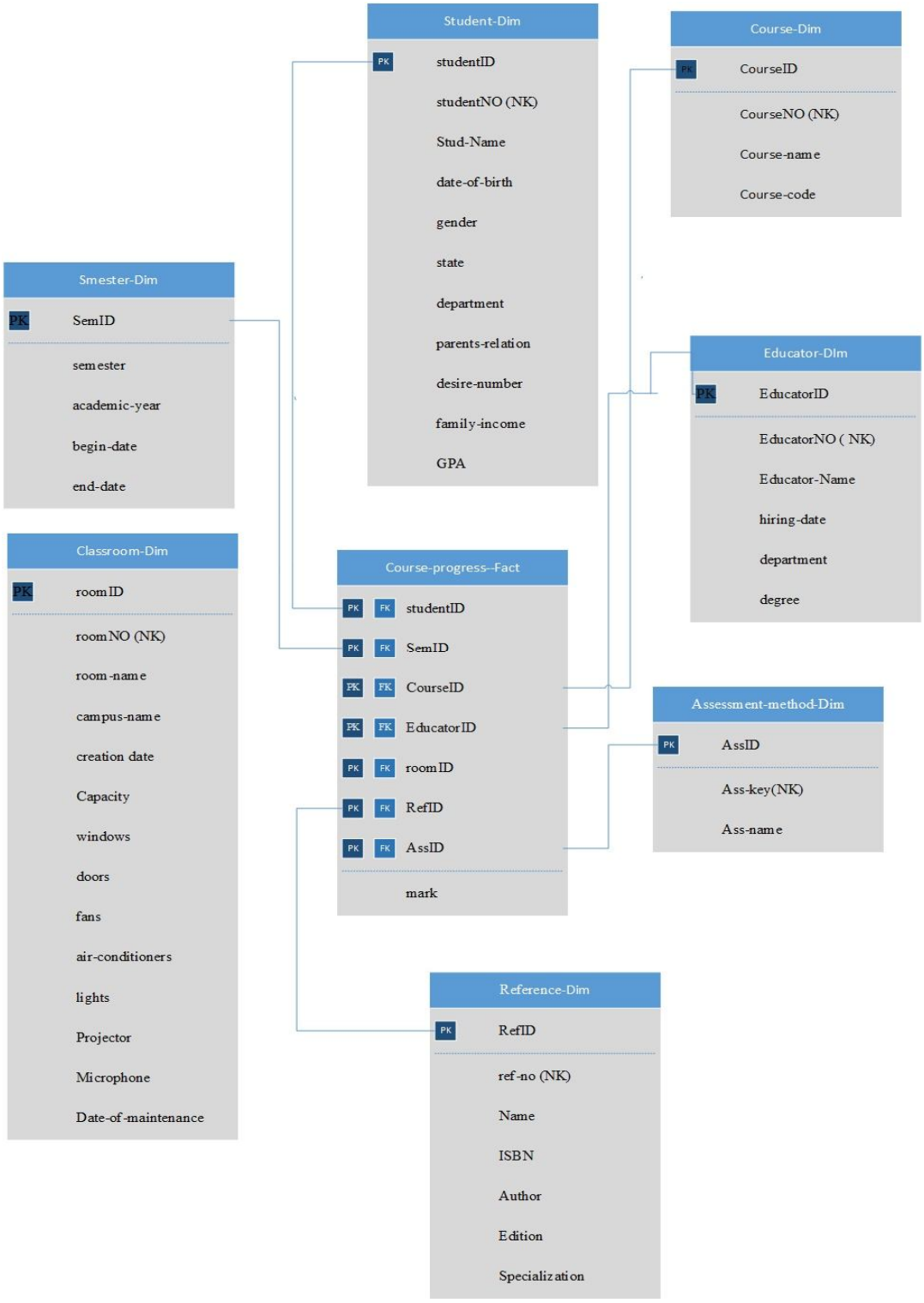


Figure 6-5 Course progress data mart

One of the most important key indicators of a successful learning process is students' academic performance. So, building a data mart to analyze data related to students' and their grades is extremely helpful in tracking the *Teaching* factor. The fact table for this data mart is called Course-Progress as shown in Figure 6-5.

The dimension tables are: *Semester* that is composed of semester id, academic year, semester begin and end dates, etc. The *Course* dimension table contains data about courses such as Course id, name, code, etc., while the *Educator* dimension includes educator id, name, qualifications, hiring date, department, etc. *Student* dimension includes student's information such as student id, name, gender, date of birth, etc. *Reference* dimension table includes reference id, ISBN, Name, etc. *Assessment method*dimension table contains data about which assessment method is used. The last dimension table is *Classroom* and it has all the information about the classroom environment including class or lecture room id, name, campus, capacity, windows, doors, light, fans, air conditioner, projector or not, and microphone or not, etc.

6.1.3.2 Class-Occupancy

Managing class and lecture rooms is one of the big challenges facing HEIs, and has a huge impact on the learning process. It is quite helpful to examine which classrooms were used the most, and least, per academic year, and to explore which had the most, and least, equipment failures, etc. This information is vital for maintenance and load management as it relates to the *Environment* factor.

The fact table in this data mart, shown in Fig 6-6, just views the relationship between all its dimensions. This type of table is called *Factless* table, as it does not contain measures(Kimball and Ross 2011). The dimension tables are: *Semester* and *Classroom*, and they are similar to the ones with the same names described in the previous Course-Progress data mart shown Fig.6-5. The *Date* dimension contains information about important dates in each semester such as date id, date, day of week such as Sunday, Monday, etc. The *Period* dimension table contains the time during the day, i.e., start and end times, for example 8 to 10, etc. The *Department* dimension includes, department id, name, college, etc. The *Occupancy* dimension contains data about occupancy status. The fact table for this data mart is *Classroom* and it consists of the keys of all the discussed dimension tables.

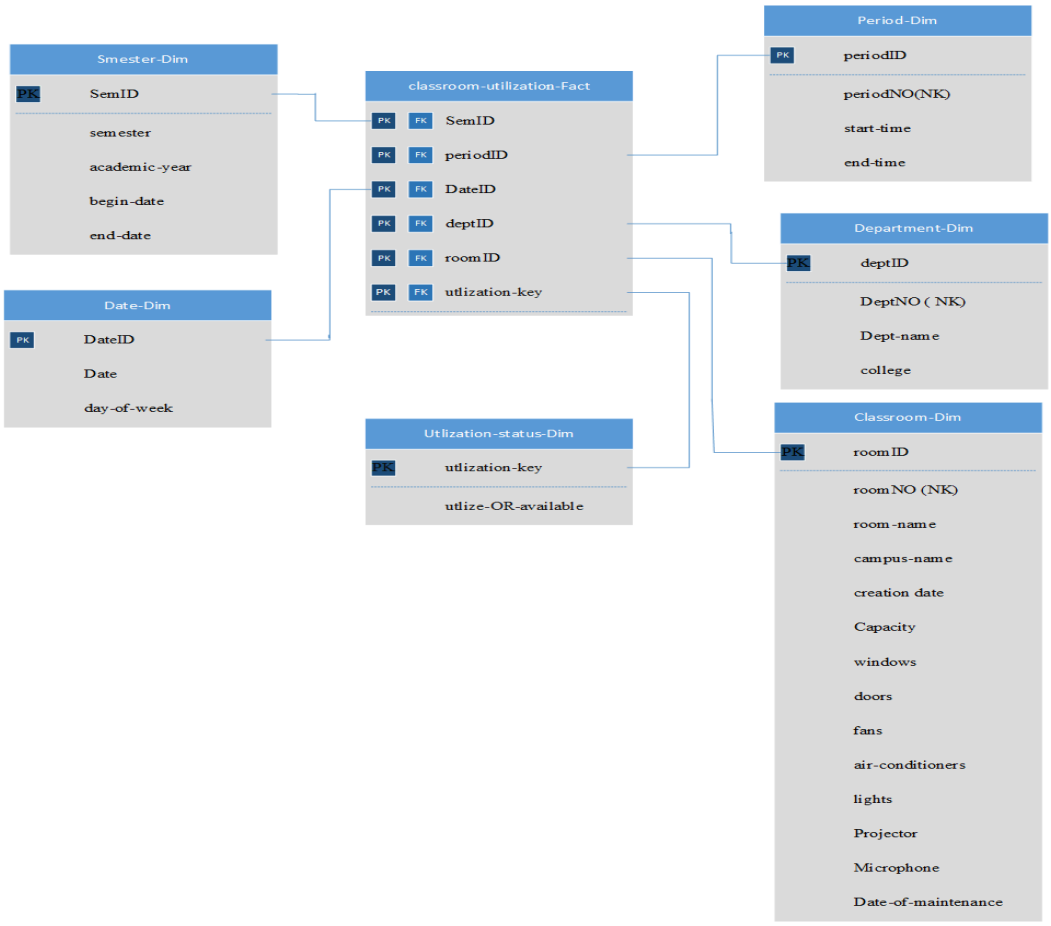


Figure 6-6 Class-Occupancy

6.1.3.3 Financial Support

This data mart is about the financial support process, here the dimensions tables are: department, student, semester, support dimension contain data about support name, support reason and period of support and sponsor dimension which contain data about sponsor and the Financial support fact table which contains keys of the above dimensional table and the value of support, Figure 6-7 present the model.

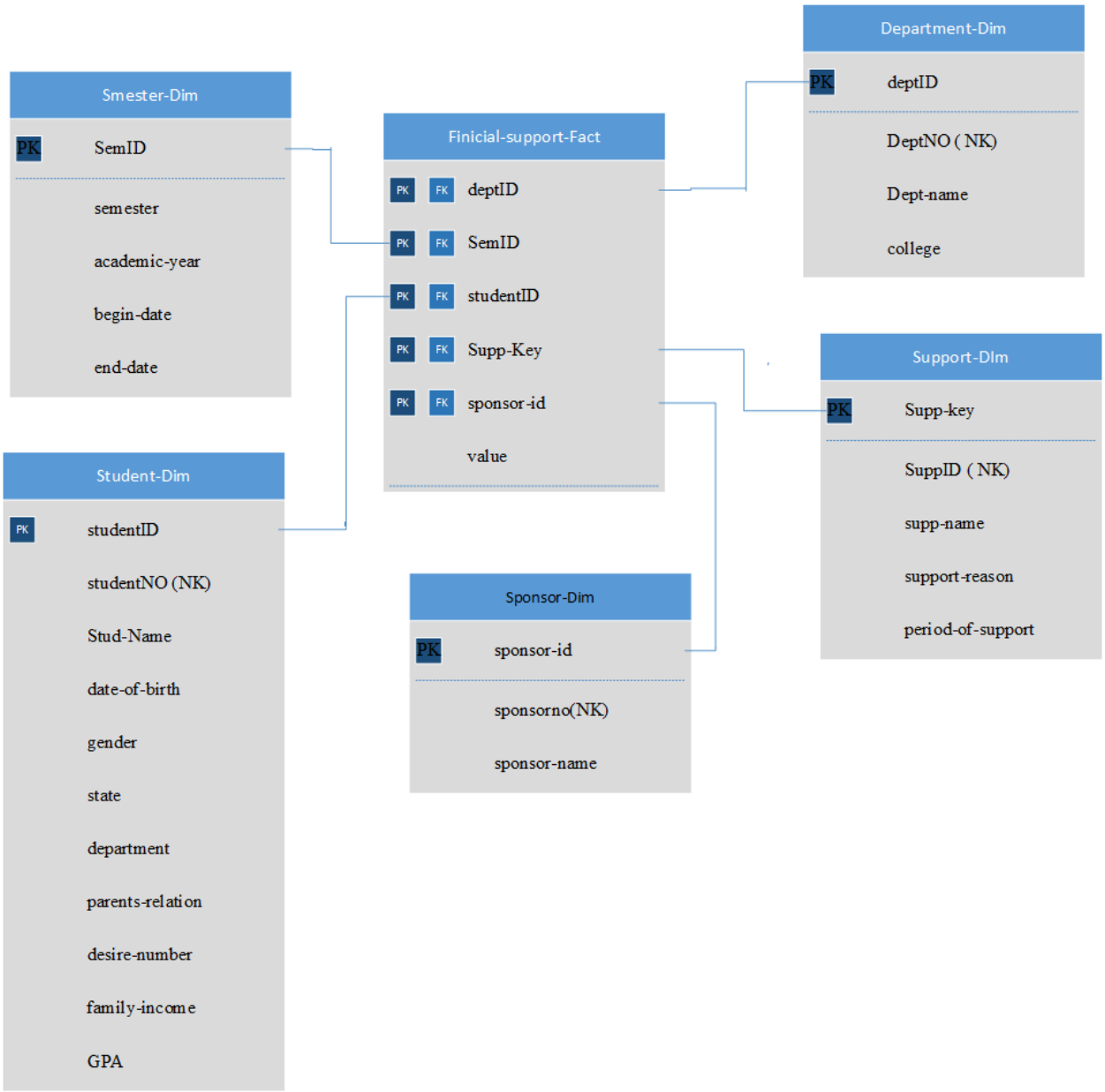


Figure 6-7 Financial Support

6.1.3.4 Students-Registration

Trends of students' enrollment numbers at the different colleges and departments over time are a strong indicator of the HEIs' competitiveness and success. This information is captured with the *HEI* and *Student* learning process factors discussed earlier. Thus, the BI model tracks students' registration and enrollment information per

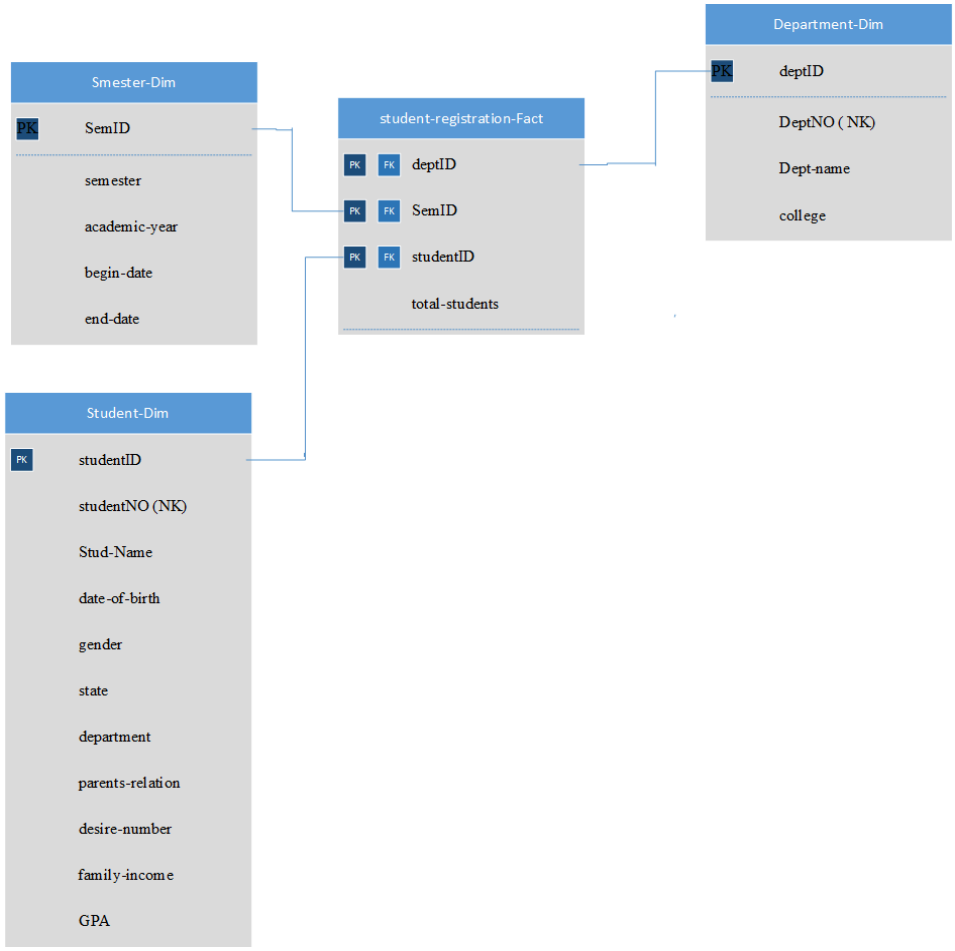


Figure 6-8 Students-Registration

department over time to enable analyzing trends.

The fact table for this data mart is Student-Registration, shown in Figure 6-8, and it contains keys of all dimension tables together with students' aggregates as measures. The dimension tables are *Semester*, *Department*, and *Student*. They are all identical to the dimension tables with the same names described in the previous data marts.

6.1.4 Data Analysis and validation

After completing the DW design, ETL procedures were carried, and the DW ready for the analytical phase. The model is implemented to explore the benefits achieved by stakeholders from adopting BI technologies and their impact on the learning process at HEIs. The open source BI tool, SpagoBI, was utilized for this phase of model implementation (Gioia, Cazzin et al. 2008).

Information dashboards and reports were created using SpagoBI that reflect key stakeholders' perspectives. The knowledge extracted from the three data marts is discussed in detail, and its impact on the study objective is outlined.

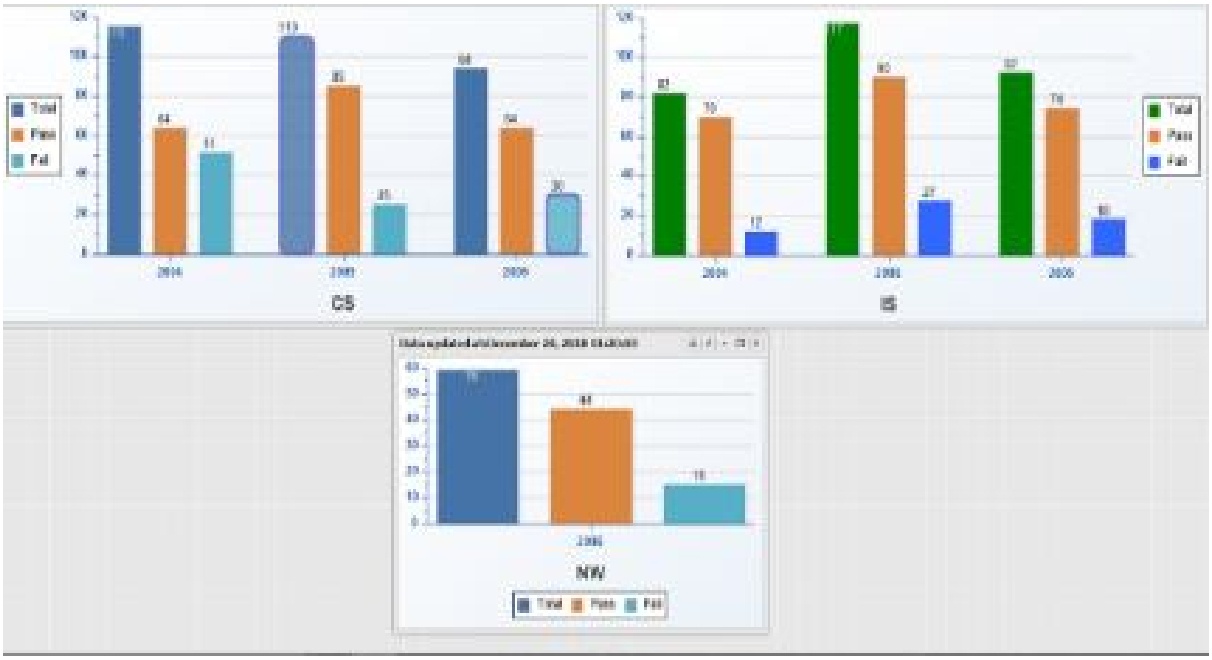


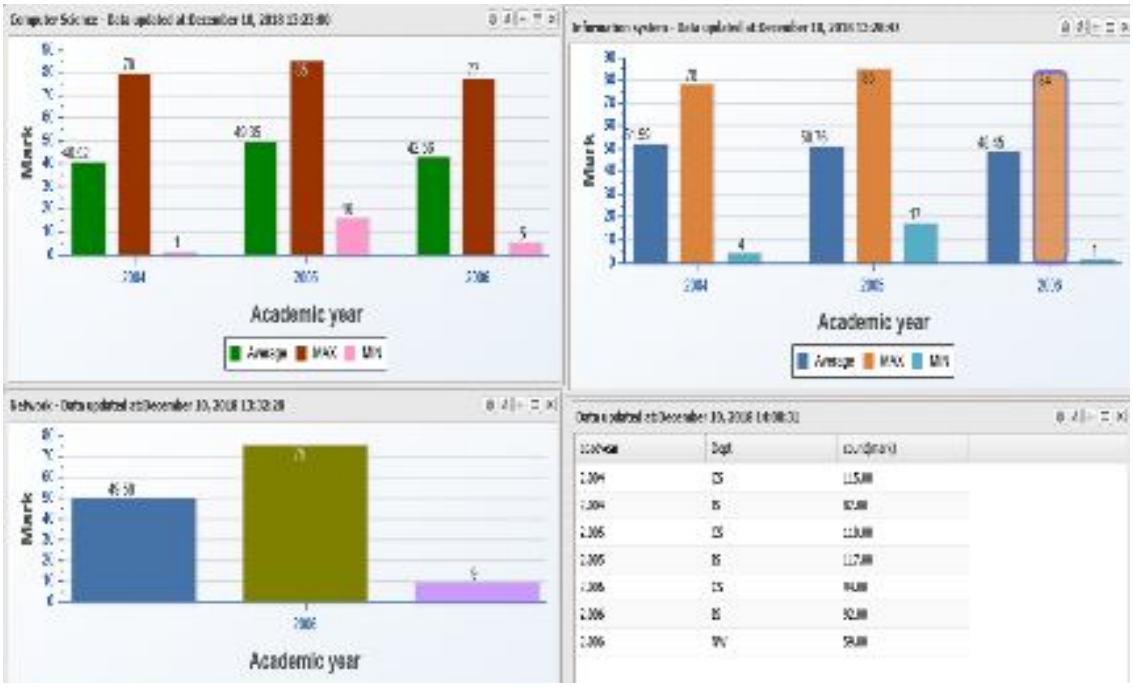
Figure 6-9 Students' performance

First, the Course-Progress data mart, avails students' performance data per course over time to key stakeholders such as college deans, department heads and educators. This academic performance information includes the total average, minimum, maximum, passing and fail students.

The dashboard shown in Figure 6-9 contains information about the *Data Structures* course taught at the College of Computer Science and Information Technology. The course was taught at the three departments in the college, namely, Computer Science (CS), Information System (IS), and Networks (NW) during the 2004, 2005, and 2006

academic years in semester four. Enrollments in the NW department began in 2005, so there is data for the 2006 academic year only. The dashboard displays the total number of students registered for the course per department together with the number of passes and fails. The percentage of student who passed the course at the CS department was 77%, 55%, and 68% for the 2004, 2005, 2006 academic years respectively. Similarly, at the IS department the results were 85%, 78%, and 80% for the same period, and 74% at the NW department in 2006. It is clear that the number of students who passed the course at the IS department is greater than that of the two other departments. This information suggests that maybe there is a difference in the teaching environment, educator, assessment methods, students, etc., and that further investigation of this finding is warranted.

Additionally, from Figure 6-9 we can see there was high failure among the CS department, 51 students, during the 2004 academic year. These finding lead the Dean and department heads to explore what made performance differ in departments while the course content is same. Correlating students' grades, educators, and assessment methods and any other information related to the *Teaching* factor might lead to some



answers. The BI model fully supports and empowers these kinds of investigations.

Figure 6-10 shows a dashboard that was created from the Course-Progress data mart. It depicts the average, maximum and minimum grade for each department during the 2004, 2005, and 2006 academic years. The dashboard clearly shows that the average grades at the CS department for this period were 40%, 49%, 42% respectively. While they were 51%, 50%, 48% at the IS department for the same period, and 49% at the NW department in 2006. The fact that the average grade was consistently less than 50% strongly indicates that the overall students' performance was not good enough, and there is weakness in *one* or *more* factor that need improvement.

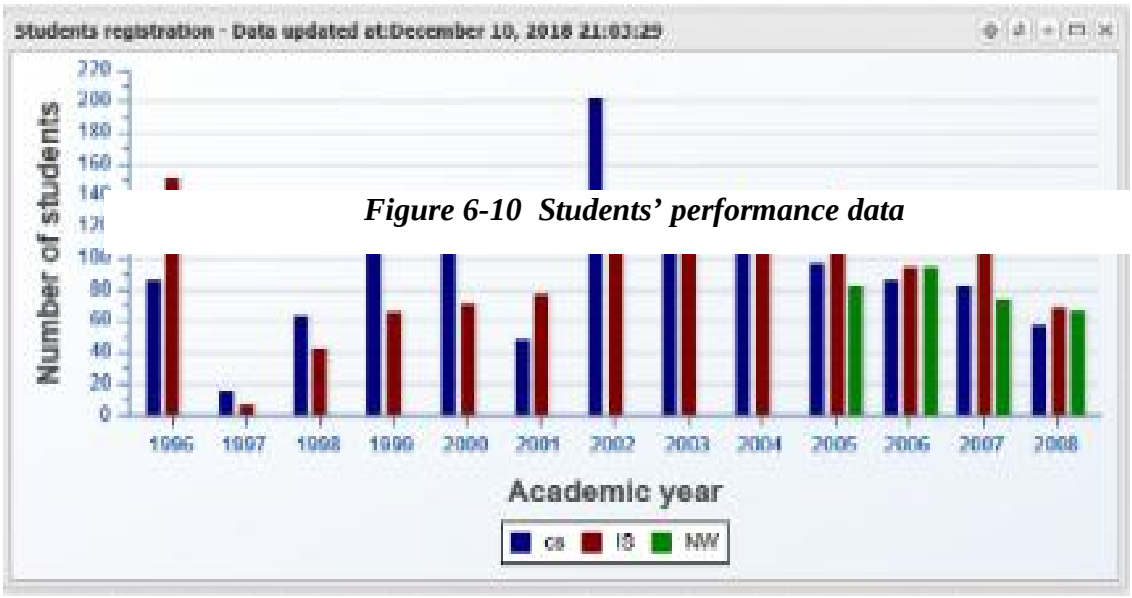


Figure 6-10 Students' performance data

Figure 6-11 Number of register first year students

The third dashboard results, shown in Figure 6-11, are from the Students-Registration data mart. It shows students' registration information for CS and IS departments between 1996 and 2008, and for the NW department from 2005 to 2008. These results show that in 1997 less than 20 students registered for both the CS and IS departments. Similarly, in 2002 the total number of students who registered for the CS department was about 200 and this is the largest number throughout the study period. The results also show that between years 2005 to 2008 the total number of registered students in the IS department was consistently greater than the CS department. All of the information derived from this dashboard poses many questions that need answers. For example, why the number of CS and IS students in 1997 was so low? What is the reason behind the IS department accepting more students than the CS and NW

departments during the 2005 to 2008 period? etc. The strength of the BI model is that it empowers such questions and inquiries while providing the tools and methods to analyze and find possible answers.

From the analysis above it is clear that using BI to investigate the many factors affecting the learning process, leads to information that can be used to determine weaknesses and to make improvement decisions.

CHAPTER 7

7 Conclusion

This chapter summarizes the approach of the thesis, reviews the contributions and discusses future work and further research possibilities.

7.1 Conclusion

This thesis investigated the adoption of BI technologies to enhance the learning process at HEIs. It developed a basic BI implementation model, identified essential factors for its adoption, and empirically examined its efficiency. The case study implementation utilized as part of the validation, focused on four of the core factors identified; environment, teaching, student and HEI.

The analysis of the results obtained from the model has demonstrated the efficiency and benefit of using BI tools to analyze data and obtain valuable decision-making information. The implemented model is very robust, the limitations is data. Additional data will lead to more information about factors, and the relationship between factors and learning process.

Furthermore, the study has laid out the foundation for improving the learning process while providing insight to HE planners, educators, and administrators to better realize the value of BI, the possible obstacles, and the existing leverage in its adoption.

7.2 Contribution and future work

Findings have demonstrated meeting the research objectives, the contribution of this research is design of Easy-to-implement BI framework which effectively serving organizations as well as students based on Scalability of the models, Strong basis for integration, analysis and cleaning administrative and learning data, and Opportunity to exploit the effect of historical data toward shaping educational and organizational goals for various educational providers and stakeholders. The scalability of the model based on easily adding new factors to extend the data models and possible BI applications in the HE domain and easily create new data marts to cover further business process in HEI.

Future works involve extending the empirical study by involving additional data and factors as well as including HEIs at other regions. Using different data analysis

techniques (data mining) discovering new interrelationships between business, educational and learning processes.

References

Abante, M. E. R., et al. (2014). "Learning styles and factors affecting the learning of general engineering students." International Journal of Academic Research in Progressive Education and Development**3**(1): 16-27.

Abramson, I. (2010). "Data Warehouse:The Choice of Inmon versus Kimball." from https://www.ismll.uni-hildesheim.de/lehre/bi-10s/script/Inmon-vs-Kimball.pdf.

Al-Rammahi, A. H. I. (2015). "Study of Simulation for Data Webhousing System by Challenging Technology and Performing " IOSR Journal of Computer Engineering.

Ali, M. M. (2013). "Role of data mining in education sector." International Journal of Computer Science and Mobile Computing**2**(4): 374-383.

Allen, H. (2017). "Why is Higher Education Important?". Retrieved 24/1, 2019, from https://www.crosswalk.com/family/homeschool/why-is-higher-education-important-1367463.html.

Amirul, N. J., et al. (2013). The physical classroom learning environment. Proceedings of the International Higher Education Teaching and Learning Conference.

Asemi, A., et al. (2011). "The role of management information system (MIS) and Decision support system (DSS) for manager's decision making process." International Journal of Business and Management**6**(7): 164-173.

Aslam, H. D., et al. (2012). "Analysing factors affecting learning environment of universities in Pakistan: A case of public sector university of Pakistan." Journal of American Science**8**(9): 58-65.

Aslam, H. D., et al. (2012). "Analyzing factors affecting students' satisfaction regarding semester system in universities of Pakistan." Journal of American Science**8**(10): 163-170.

Aziz, A. A., et al. (2014). "A FRAMEWORK FOR EDUCATIONAL DATA WAREHOUSE (EDW) ARCHITECTURE USING BUSINESS INTELLIGENCE (BI) TECHNOLOGIES." Journal of Theoretical & Applied Information Technology**69**(1).

Aziz, A. A., et al. (2012). "Intelligent System for Personalizing Students' Academic Behaviors-A Conceptual Framework." International Journal on New Computer Architectures and Their Applications**2**(1): 138-153.

Baradwaj, B. K. and S. Pal (2012). "Mining educational data to analyze students' performance." arXiv preprint arXiv:1201.3417.

Bates, E. (2015). "UVM Big Data? Aggregating Campus Databases and Creating a Data Warehouse to Improve Student Retention Rates at the University of Vermont."

Berkhin, P. (2006). A survey of clustering data mining techniques. Grouping multidimensional data, Springer**:** 25-71.

Betoret, F. D. and A. D. Tomás (2003). "Evaluation of the university teaching/learning process for the improvement of quality in higher education." Assessment & evaluation in higher education**28**(2): 165-178.

Bhise, R., et al. (2013). "Importance of data mining in higher education system." IOSR Journal Of Humanities And Social Science (IOSR-JHSS)**6**(6): 18-21.

Borkar, S. and K. Rajeswari (2013). "Predicting students academic performance using education data mining." International Journal of Computer Science and Mobile Computing**2**(7): 273-279.

Bouckaert, R. R., et al. (2010). "WEKAâˆ'Experiences with a Java Open-Source Project." Journal of Machine Learning Research**11**(Sep): 2533-2541.

Brown, E. D. (2015). from https://larion.com/resource/whats-difference-business-intelligence-big-data/components-of-business-intelligence/.

Chan, R. Y. (2016). "Understanding the purpose of higher education: An analysis of the economic and social benefits for completing a college degree." Journal of Education Policy, Planning and Administration**6**(5): 1-40.

Cook, S. (1999). "A chronology of academic advising in America." The Mentor: An Academic Advising Journal**1**(2).

Dell'Aquila, C., et al. (2008). "Business intelligence applications for university decision makers." WSEAS Transactions on Computers**7**(7): 1010-1019.

Durairaj, M. and C. Vijitha (2014). "Educational data mining for prediction of student performance using clustering algorithms." International Journal of Computer Science and Information Technologies**5**(4): 5987-5991.

EDM.org (2011). "Educational Data Mining ". from https://educationaldatamining.org/.

Education, T. F. o. H. (2000). Higher education in developing countries: Peril and promise, World Bank.

Falakmasir, M. H., et al. (2010). Business intelligence in e-learning:(case study on the Iran university of science and technology dataset). Software Engineering and Data Mining (SEDM), 2010 2nd International Conference on, IEEE.

Gioia, A., et al. (2008). SpagoBI: A distinctive approach in open source business intelligence. 2008 2nd IEEE International Conference on Digital Ecosystems and Technologies, IEEE.

Gulliver, J. H. (2014). "Top 4 Benefits of Higher Education." Retrieved 21/3, 2019, from http://www.goodchoicesgoodlife.org/choices-for-young-people/the-benefits-of-higher-education/.

Han, J., et al. (2011). Data mining: concepts and techniques, Elsevier.

Harilatha, U. and N. Sudhakaryadav (2014). "Predicting Educational Performance of a Student Failure and Dropout by using Data mining Techniques." IJCSIT) International Journal of Computer Science and Information Technologies**5**(6).

Hema, R. and N. Malik (2010). Data mining and business intelligence. Proceedings of the 4th National Conference.

Inmon, W. H. (2005). Building the data warehouse, John wiley & sons.

JICA (2004). "Approaches for Systematic Planning

of Development Projects/higher education."

Julaily Aida, J., et al. (2012). "Conceptual Framework of Business Intelligence Analysis Academic Environment using BIRT."

Kabakchieva, D. (2013). "Predicting student performance by using data mining methods for classification." Cybernetics and Information Technologies**13**(1): 61-72.

Kabakchieva, D. (2015). "Business intelligence systems for analyzing university students data." Cybernetics and Information Technologies**15**(1): 104-115.

Kelly, D. (2005). "Business Intelligence: the smart way to track academic collections." Educause Quarterly**28**(4): 48-53.

Khan, M. A., et al. (2014). Data mining techniques for business intelligence in educational system: A case mining. 2014 World Congress on Computer Applications and Information Systems (WCCAIS), IEEE.

Khan, R. A. and S. Quadri (2012). "Business intelligence: an integrated approach." Business Intelligence Journal**5**(1): 64-70.

Khedr, A. E. and A. I. El Seddawy (2015). "A proposed data mining framework for higher education system." International Journal of Computer Applications**113**(7).

Kimball, R. and M. Ross (2011). The data warehouse toolkit: the complete guide to dimensional modeling, John Wiley & Sons.

Kumar, D. V., J. Senthil, and Ramesh Palanisami "Optimizing Resources in Teaching and Learning using Cloud Technology."

Kumar, V. and A. Chadha (2011). "An empirical study of the applications of data mining techniques in higher education." International Journal of Advanced Computer Science and Applications**2**(3).

Lachman, S. J. (1997). "Learning is a process: Toward an improved definition of learning." The Journal of psychology**131**(5): 477-480.

Lau, L. K. (2003). "Institutional factors affecting student retention." Education-Indianapolis then Chula Vista-**124**(1): 126-136.

Leung, S.-O. (2011). "A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales." Journal of Social Service Research**37**(4): 412-421.

Michigan, s. (2017). "Difference-between-business-intelligence-vs-business-analytics." from https://www.michiganstateuniversityonline.com/resources/business-analytics/difference-between-business-intelligence-vs-business-analytics/.

Mohd, W. M. B. W., et al. (2008). A knowledge-based digital dashboard for higher learning institutions. Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer.

Nikky (2010). "Teaching Aids, Their Needs, Types and Importance Of Teaching Aids In Teaching Learning Process." Retrieved 28/3, 2019, from https://www.indiastudychannel.com/resources/120148-Teaching-Aids-Their-Needs-Types-and-Importance-Of-Teaching-Aids-In-Teaching-Learning-Process.aspx.

Olson, D. L. and D. Delen (2008). Advanced data mining techniques, Springer Science & Business Media.

Oracle, S. "Developer-Data Modeler User's Guide." E15802-03.

ORACLE®. "Introduction to Data Warehousing Concepts." Retrieved 16/12, 2020, from https://docs.oracle.com/database/121/DWHSG/concept.htm#DWHSG-GUID-452FBA23-6976-4590-AA41-1369647AD14D.

Pal, S. (2012). "Mining educational data to reduce dropout rates of engineering students." International Journal of Information Engineering and Electronic Business**4**(2): 1.

Pavione, C. S. S. N., et al. (2016). "Factors that Influence the Teaching-Learning Process from the Perspective of Accountancy Students: Analysis at a Higher Education Institution in Minas Gerais." Revista de Educação e Pesquisa em Contabilidade**10**(2).

Piedade, M. B. and M. Y. Santos (2009). Business Intelligence supporting the teaching-learning process. WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering, Citeseer.

Piedade, M. B. and M. Y. Santos (2010). Promoting Students Success with a Business Intelligence System. CSEDU (2).

Pomerol, J.-C. and F. Adam (2004). Practical decision making–From the legacy of Herbert Simon to decision support systems. Actes de la Conférence Internationale IFIP TC8/WG8.

Power, D. J. (2007). "A brief history of decision support systems." DSSResources. com**3**.

Puteh, M., et al. (2015). "The classroom physical environment and its relation to teaching and learning comfort level." International Journal of Social Science and Humanity**5**(3): 237-240.

Rangra, K. and K. Bansal (2014). "Comparative study of data mining tools." International journal of advanced research in computer science and software engineering**4**(6).

Ranjan, J. (2009). "Business intelligence: Concepts, components, techniques and benefits." Journal of Theoretical and Applied Information Technology**9**(1): 60-70.

Romero, C. and S. Ventura (2010). "Educational data mining: a review of the state of the art." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**40**(6): 601-618.

Rouhani, S., et al. (2012). "Review study: business intelligence concepts and approaches." American Journal of Scientific Research**50**(1): 62-75.

Saleh, A. R., et al. (2008). "Academic business intelligence system development using SAS tools."

Simon, H. A. (1997). "Behavior]: Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations. 4." Aufl., New York ua.

Singh, R., et al. (2013). "An Empirical Study of Applications of Data Mining Techniques for Predicting Student Performance in Higher Education." International Journal of Computer Science and Mobile Computing**2**: 53-57.

Solomon, N. and G. Paul (2004). "Business intelligence." Communications of the Association for Information Systems**13**: 450.

Stair, R. and G. Reynolds (2013). Principles of information systems, Cengage Learning.

Teal, F. (2011). "Higher education and economic development in Africa: A review of channels and interactions." Journal of African Economies**20**(suppl_3): iii50-iii79.

Turban, E. (2011). Decision support and business intelligence systems, Pearson Education India.

Turban, E., et al. (2014). Business intelligence and analytics: systems for decision support, Pearson Higher Ed.

Watson, H. J. (2009). "Tutorial: Business intelligence-Past, present, and future." Communications of the Association for Information Systems**25**(1): 39.

Yadav, S. K. and S. Pal (2012). "Data mining application in enrollment management: A case study." International Journal of Computer Applications**41**(5).

Young-Jones, A. D., et al. (2013). "Academic advising: does it really impact student success?" Quality Assurance in Education**21**(1): 7-19.

Zhang, Y., et al. (2010). "Using data mining to improve student retention in HE: a case study."

Zilli, D. (2014). Self-service business intelligence for higher education management. Management, Knowledge and Learning, International Conference, Portoroź, Slovenia. Retrieved from http://www. toknowpress. net.

Appendix-A Students Questionnaire

Gender : Male ☐ Female ☐	**Year :** 1st ☐ 2st ☐ 3st ☐ 4st ☐
College:	**Department:**

#	***Statements***	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
	Environment					
1.	Roads well paved inside the university					
2.	campuses Adjacent and there is no problem in mobility and movement between them					
3.	Public transportation available to get to College and no difficulty in its availability					
4.	Campus buildings and terrain are well maintained.					
5.	The lectures halls are air conditioned and good ventilation					
6.	There is no external noise effect inside the lectures halls					
7.	Lighting in the lectures halls is good					
8.	The sound inside the halls hears clearly					
9.	The number of students within the Hall is suitable					
10	The seats inside the halls are comfortable					
11	laboratories well designed and equipment are available					
12	The library is well equipped and comfortable					
13	The library contains modern references in the field of specialization					
14	There is a rooms in the library for group work (so you do not disturb others)					
15	The University offers adequate health service and emergency support.					
16	The cafeterias services is good					
17	The price of food is acceptable					
18	All rooms and facilities accessible by disabled people /wheel chair drivers					
	Teaching					
19	Lectures schedule is set up properly					

20	There is a time in the schedule to do other activities such as (praying, eating ...)					
21	I don't have a problem with the language (English) used in teaching					
22	Teacher teaching method suits my way of learning					
23	Teaching is done in the traditional way (give information by teacher only)					
24	Teacher provides a suitable learning environment allows discussion and debate and exchange of views					
25	Teacher encourages teamwork and collaborative environment					
26	Teachers' knowledge and teaching methods are up to date in each course.					
27	The teacher uses aids in teaching such as (projector, computer, pictures, models, any other ways)					
28	The teacher uses examples of real life applications to support theoretical foundations.					
29	Learning material is published in digital as well as printed form					
30	There is a suitable number of available references for study material to follow the courses					
31	The students are divided into groups to perform projects and assignments with the ability to organize and communicate between members of the Group					
32	Information is displayed in orderly and clearly form					
33	There's a chance to communicate with the teacher out of the class room					
34	The teacher uses social media to communicate with students					
35	The teacher uses a number of methods such as (tests, seminars, midterm exam) in assess student performance					
36	Exams questions are clear and cover					

	the content of the course.					
37	There is diversity of questions (subjective, objective, analysis) in exams and tests					
38	There is practical semester (in a real working environment) in the curriculum					
39	The office hours are adequate					
	academic advising					
40	Academic Advisor offers a suitable number office hours to meet students					
41	I can suitably communicate with the academic advisor					
42	Academic Advisor suitably helps me in determining my goals					
	Student					
43	The study in this college was my desire and decision.					
44	I can communicate and collaborate with colleagues					
45	there are groups and teams where i can use my knowledge and exchange/interact with others					
46	The relationship with my parents is stable					
47	I can communicate with my parents and reflect problems that I am suffering from					
48	I can pay tuition fees without problems					
49	University allows time for diverse activities (political, sports, cultural, …)					
50	The University offers suitable places and infrastructure to practice these activities					
51	I understand the rules and regulations of the academic system					
52	I can communicate with persons who are responsible for student affairs					
53	i am able to work along with the studies and get practical experience					
	HEI					
54	Curriculum design and approval processes involve industrial and business experts and students.					

55	The curriculum hasa suitable sequence for the students.					
56	The University provides vocational training opportunities as well as postgraduate studies.					
57	University has international research projects and integrates results in teaching.					
58	University has research projects with various sectors in country (industry, health…) and integrates results in teaching.					
59	The University provides financial support to poor students					
60	University provides incentives for superior students					
61	University hasoffice dedicated to student affairs					
62	University has office dedicated to examination					
63	When I have a problem I know which council is responsible for solving it					
	ICT support					
64	Enough Computers available					
65	University provides technical support when using IT					
66	University provides fast internet					
67	University offers learning management system (LMS)					
68	I am using the LMS					
69	There is IT supported collaborative work environment					
70	The use of personal data is securely treated.					
71	There is Wi-Fi access all over the campus					

Appendix-B Questionnaire analysis

Data analysis in details using WEKA

A. Environment

a. University infrastructure dimensions

From table (1) of results we can see that the responses of students to university infrastructure dimensions as follow:

- Roads well paved dimension, 41% of students are disagree, 30% are strongly disagreed and 29% are agreed. This mean 71% are disagreed.
- Mobility and movement dimension, 71% of students are disagreeing and 29% are agreed.
- Building Maintenance, 71% of students are disagreed and 29% are disagreed.
- Facilities for disabled people, all students are strongly disagreed.
- Public Transportation, 71% of students are strongly disagreeing to this dimension and 29% are disagreed.

Table1 : university infrastructure dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		41%	30%	29%
1	Roads well paved inside the university	2.9	2.6	3.7
2	Campuses Adjacent and no problem in mobility and movement between them	3.4	2.9	3.8
3	Campus buildings and terrain are well maintained	2.4	2.3	2.9
4	All rooms and facilities accessible by disabled people /wheel chair drivers	2.2	2.3	2.5
5	Public transportation available to College and no difficulty in its availability	2.4	2.3	2.8

b. Classroom environment dimensions

From table (2) of results we can see that the responses of students to classroom environment dimensions as follow:

- With regard to air condition and ventilation, 70% of student agreed and 30% strongly disagreed that the air condition and ventilation is good.
- All students seen there is external noise effect inside lectures hall.
- 70% of students agreed that the lighting is good and 30% are strongly disagreed.
- Sound hears Cleary, 41% disagreed, 30% strongly disagreed and 29% are agreed. This means 71% are disagreed.
- Number of students, 70% agreed that the number of students is suitable and 30% are disagreed.
- Seats are comfortable, 71% are strongly disagreed and 29% are agreed.

Table02: classroom environment dimension

	Dimensions	**Clusters and students percentage**		
		cluster0	cluster1	cluster2
		41%	30%	29%
1	The lectures halls are air conditioned and good ventilation	4.0	2.0	4.1
2	There is no external noise effect inside the lectures halls	3.0	2.5	3.4
3	Lighting in the lectures halls is good	4.2	2.6	4.5
4	The sound inside the halls hears clearly	3.1	2.1	3.9
5	The number of students within the Hall is suitable	3.8	3.4	4.5
6	The seats inside the halls are comfortable	2.4	2.6	4.0

c. Study environment dimensions

Regarding to study environment dimensions table-3 presents the results of students' responses to study environment as follow:

- 71% of students are strongly disagreed about laboratories design and equipment and 29% are agreed.
- 41% of students are disagreed that the library is well design, 30% strongly disagreed and 29% are strongly greed.

- About references 71% strongly disagreed that the library contain modern references and 29% are agreed.
- With regard to existence of rooms for group work in library all students are strongly disagreed.

Table03: Study environment dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		41%	30%	29%
1	laboratories well designed and equipment are available	2.4	2.3	3.8
2	The library is well equipped and comfortable	3.6	2.7	4.6
3	The library contains modern references in the field of specialization	2.4	2.1	4.1
4	There is a rooms in the library for group work (so you do not disturb others)	2.3	2.2	2.6

d. University services dimensions

From table 4 of results we can see that:

- 71% of students are strongly disagreed that university offer adequate health services and 29% of students are disagreed.
- About cafeterias services, 70% of students are disagreed that services is good and 30% are strongly disagrees.
- Price of food is acceptable, all students are strongly disagreed.

Table04: University services dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		41%	30%	29%
1	The University offers adequate health service and emergency support.	2.1	2.1	3.0
2	The cafeterias services is good	2.8	2.3	3.0
3	The price of food is acceptable	2.5	1.9	2.7

B. Teaching

a. Timetable and language barrier dimensions

According to results from table 5:

- 59% of students are strongly disagreed and 41% of students disagreed that the schedule is set up properly.
- 59% of students are strongly disagreed and 41% of students disagreed that there is time in schedule to do other activities.
- I don't have problem with teaching language, 63% of students are disagreed and 37% are strongly disagreed.

Table05: Timetabling and language barrier

Dimensions		**Clusters and students percentage**		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	Lectures schedule is set up properly	2.3	2.9	2.1
2	There is a time in the schedule to do other activities such as (praying, eating ...)	2.3	2.9	2.3
3	I don't have a problem with the language (English) used in teaching	2.9	3.3	2.4

b. Teaching Method dimensions

Table 6 present results of teaching method dimensions:

- 59% of students arestrongly disagreed with teaching methods and 41% are disagreed.
- The second dimension is about teaching done in traditional way, all responses around strongly disagreed.
- About teacher's knowledge and teaching methods, 59% of students are strongly disagreed and 41% of students are disagreed.
- 22% of students disagree with information displayed in orderly and clearly form,37% are strongly disagreed and 41% are agreed.

Table06: Teaching method dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	Teacher teaching method suits my way of learning	2.2	2.9	2.1
2	Teaching is done in the traditional way (give information by teacher only)	2.7	2.7	2.6
3	Teachers' knowledge and teaching methods are up to date in each course.	2.4	2.8	2.2
4	Information is displayed in orderly and clearly form	2.8	3.9	1.9

c. *Teaching support dimensions*

The result in table 7 is about the teaching support dimensions; from table we can see that

- 63% of students agree with teacher uses aids and 37% disagreed.
- 22% agreed that teacher use real life examples, 41% are disagree and 37% strongly disagreed.
- 63% are agreed the materials published in digital and printed form and 37% are disagreed.
- 22% are neutral, 41% are agreed and 37% are strongly disagreed.

Table07: Teaching support dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	The teacher uses aids in teaching such as (projector, computer, pictures, models, any other ways)	4.1	4.3	3.4
2	The teacher uses examples of real life applications to support theoretical foundations.	3.9	3.6	2.3
3	Learning material is published in digital as well as printed form	3.7	4.3	3.0
4	There is a suitable number of available references for study material to follow the courses	1.4	4.2	2.5

d. Communication dimensions

Table 8presents the communication dimensions results as follow:

- All students strongly disagreed with office hours.
- 63% of students disagreed that there is chance to communicate with teachers out of the classroom and 37% are strongly disagreed.
- 63% of students disagreed that the teachers' uses social media in communication and 37% are strongly disagreed.

Table08: Communication dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	The office hours are adequate	2.5	2.7	2.0
2	There's a chance to communicate with the teacher out of the class room	3.1	3.1	2.3
3	The teacher uses social media to communicate with students	3.0	3.0	2.1

e. *Support of collaborative work dimensions*

Results in table 9 show that:

- 63% of disagreed that teacher encourages teamwork and collaborative environment and 37% are strongly disagreed.
- About dividing students to group, 59% of students are disagreed and 41% are agreed.
- Teacher provides a suitable learning environment, 63% of students are disagreed and 37% are strongly disagreed.

Table09: Support of collaborative work

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	Teacher encourages teamwork and collaborative environment	3.6	3.4	2.1
2	The students are divided into groups to perform projects and assignments with the ability to organize and communicate between members of the Group	3.6	4.0	2.8
3	Teacher provides a suitable learning environment allows discussion and debate and exchange of views	3.5	3.4	2.1

f. *Examination and assessment methods*

From table 10 we can see that:

- 63% of students were agreed that teacher uses number of methods in assess the performance and 37% were disagreed.
- Exams question are clear and cover the content, 63% of students were disagreed and 37% were strongly disagreed.
- About the diversity of exams questions, 63% of students were disagreed and 37% were strongly disagreed.

Table010: Examination and assessment methods

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		22%	41%	37%
1	The teacher uses a number of methods such as (tests, seminars, midterm exam) in assess student performance	4.3	4.3	3.4
2	Exams questions are clear and cover the content of the course.	2.8	3.1	2.2
3	There is diversity of questions (subjective, objective, analysis) in exams and tests	3.2	3.4	2.2

C. Academic advising

With regard to academic advice dimension from table 11 of results we can see that:

- 20% of students wereneutral, 30% were agreed and 50% were strongly disagreed that the office hours is suitable.
- 50% were disagree that there can suitably communicate with advisor.
- 50% were disagreed that the advisor helps to determine their goals.

Table011: Academic advice dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		20%	30%	50%
1	Academic Advisor offers a suitable number office hours to meet students	1.2	3.7	2.2
2	I can suitably communicate with the academic advisor	3.5	3.4	1.7
3	Academic Advisor suitably helps me in determining my goals	3.2	3.6	1.5

D. Student

a. Family support dimension

From table 12 of results, the students' responses to family support dimensions as follow:

- Relationship with parents are stable, 20% of students were disagreed and 80% were strongly agreed.
- Communication with parents, 20% of students were disagreed and 80% agreed.
- Paying the tuition fees, 56% of students were strongly disagreed and 44% of students were agreed that they don't have problem to pay tuition fees.

Table012: Family support dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		20%	44%	36%
1	The relationship with my parents is stable	3.6	4.8	4.8
2	I can communicate with my parents and reflect problems that I am suffering from	2.8	4.5	4.1
3	I can pay tuition fees without problems	2.5	4.1	2.7

b. Student dimensions

Table013: Student dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		20%	44%	36%
1	The study in this college was my desire and decision.	2.5	3.7	3.7
2	I can communicate and collaborate with colleagues	2.4	4.4	4.4
3	there are groups and teams where i can use my knowledge and exchange/interact with others	1.9	3.7	4.1
4	I understand the rules and regulations of the academic system	2.3	4.0	2.3
5	I can communicate with persons who are responsible for student affairs	2.2	3.1	1.9
6	i am able to work along with the studies and get practical experience	2.5	3.0	2.3

Table 13 presents the results to personality dimensions; from these results we can see that:

- Study in college is desire and decision of students, 80% agreed and 20% were strongly disagreed.
- Communication with colleague, 80% of students agreed and 20% were strongly disagreed.
- There are groups and teams to exchange knowledge, 20% were strongly disagreed and 80% were agreed.
- Understanding of rules, 56% of students were strongly disagreed and the 44% were agreed.
- Communication with student affair person, 56% were strongly disagreed and 44% weredisagreed.
- Working with study, 56% were strongly disagreed and 44% weredisagreed.

c. *Activities dimensions*

From table 14 we see that:

- 56% of students strongly disagreed that there is time and suitable places to practice activities and 44% are disagreed.

Table014: Activities dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		20%	44%	36%
1	University allows time for diverse activities (political, sports, cultural, …)	2.2	3.0	2.0
2	The University offers suitable places and infrastructure to practice these activities	2.3	2.9	2.0

E. HEI

a. *Curriculum design*

The result from table 15 shows that:

- Involving industrial, business experts and students in curriculum design processes, 27% of students were disagreed, 34% were strongly disagree and 39% were neutral.
- Sequence of curriculum is stable, 27% of students were disagreed and 73% were strongly disagreed.
- There is practical semester in curriculum; all students were strongly disagreed to thispoint.

Table015: Curriculum design dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		27%	39%	34%
1	Curriculum design and approval processes involve industrial, business experts and students.	2.8	1.8	1.9
2	The curriculum has a suitable sequence for the students.	3.1	2.7	2.1
3	There is practical semester (in a real working environment) in the curriculum	2.7	2.6	2.1

b. Additional education opportunities dimensions

From table 16 of results we can see that:

- 73% of students were strongly disagreed with university provides additional vocational training dimension and 27% were agreed.
- About the research outside and inside country, 27% of students were agreed, 34% were strongly disagreed and 39% were neutral.

Table016: Additional education opportunities dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		27%	39%	34%
1	The University provides vocational training opportunities as well as postgraduate studies.	4.0	2.5	1.9
2	University has international research projects and integrates results in teaching.	4.0	1.7	2.0
3	University has research projects with various sectors in country (industry, health…) and integrates results in teaching.	3.7	1.6	1.9

c. Support for students dimensions

Table 17 presents the results about support for students; from result we can see that:

- 27% of students were disagreed that there is support for poor students and 73% were strongly disagreed.
- 66% of students were disagreed with incentive for superior and 34% were strongly disagreed.

Table17: Support for students

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		27%	39%	34%
1	The University provides financial support to poor students	3.1	2.5	1.9
2	University provides incentives for superior students	3.6	3.4	1.9

d. Students' affairs

From table 18 of results we can see that:

- With regard to office dedicated to students affairs 66% of students were agreed and 34% were strongly disagreed that there office dedicated to students affairs.
- University has office dedicated to examination, 34% of students were neutral, 39% weredisagreed and 27% were agreed.
- When I have a problem I know which council is responsible for solving it, 27% of students were disagreed and 73% were strongly disagreed.

Table018: Students' affairs

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		27%	39%	34%
1	University has office dedicated to students affairs	4.0	4.0	2.0
2	University has office dedicated to examination	3.8	3.2	1.8
3	When I have a problem I know which council is responsible for solving it.	3.2	2.6	2.0

F. ICT

a. Technical support dimensions

From table 19 of results we can that:

- Enough computers available 41% of students were agreed and 59% were strongly disagreed.
- Providing of technical support 41% of students were agreed that university provides technical support and 59% were neutral.

Table019: Technical support dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		15%	26%	59%
1	Enough Computers available	4.1	4.1	2.4
2	University provides technical support when using IT	4.0	3.9	1.8

b. Availability of network and internet

From table 20 of results we can see that:

- University provides fast internet, 15% of students agreed, 26% were disagreed and 59% were strongly disagreed.
- The use of personal data is securely treated, 15% of students agreed and 85% were strongly disagreed.
- There is Wi-Fi access, 15% of students disagreed and 85% were strongly disagreed.

Table020: availability of network and internet dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		15%	26%	59%
1	University provides fast internet	3.8	3.1	2.3
2	The use of personal data is securely treated.	4.2	2.5	2.1
3	There is Wi-Fi access all over the campus	2.8	2.1	2.1

c. *Availability of Learning Management System (LMS)*

Table 21 shows that:

- 15% of students disagreed that university offers LMS, 26% were neutral and 59% were strongly disagreed.
- 15% of students disagreed that they use LMS and 85% were strongly disagreed.
- 15% of students agreed that there is IT support for collaborative work, 26% were neutral and 59% were strongly disagreed.

Table021: Availability of LMS dimensions

Dimensions		Clusters and students percentage		
		cluster0	cluster1	cluster2
		15%	26%	59%
1	University offers learning management system (LMS)	3.5	1.7	2.1
2	I am using the LMS	3.6	1.9	2.1
3	There is IT supported collaborative work environment	3.7	1.7	2.2

Data analysis in details using SPSS

A. Environment

Statements	Mean	Result	Std. Deviation
Roads well paved inside the university	3.02	neutral	1.2567
campuses Adjacent and there is no problem in mobility and movement between them	3.364	neutral	1.2719
Public transportation available to get to College and no difficulty in its availability	2.465	Disagree	0.9692
Campus buildings and terrain are well maintained.	2.517	Disagree	1.2043
The lectures halls are air conditioned and good ventilation	3.454	Agree	1.4037
There is no external noise effect inside the lectures halls	2.974	Neutral	1.2497
Lighting in the lectures halls is good	3.775	Agree	1.2677
The sound inside the halls hears clearly	3.046	Neutral	1.3782
The number of students within the Hall is suitable	3.864	Agree	1.2142
The seats inside the halls are comfortable	2.919	Neutral	1.3467

laboratories well designed and equipment are available	2.801	Neutral	1.3505
The library is well equipped and comfortable	3.616	Agree	1.4365
The library contains modern references in the field of specialization	2.792	Neutral	1.4872
There is a rooms in the library for group work (so you do not disturb others)	2.367	Disagree	0.9174
The University offers adequate health service and emergency support.	2.396	Disagree	1.1805
The cafeterias services is good	2.705	Neutral	1.2648
The price of food is acceptable	2.37	Disagree	1.0281
All rooms and facilities accessible by disabled people /wheel chair drivers	2.329	Disagree	1.0107

B. Teaching

Lectures schedule is set up properly	2.474	Disagree	1.1168
There is a time in the schedule to do other activities such as (praying, eating ...)	2.532	Disagree	1.0823
I don't have a problem with the language (English) used in teaching	2.876	Neutral	1.3359
Teacher teaching method suits my way of learning	2.413	Disagree	1.3142
Teaching is done in the traditional way (give information by teacher only)	2.665	Neutral	1.3374
Teacher provides a suitable learning environment allows discussion and debate and exchange of views	2.954	Neutral	1.4093
Teacher encourages teamwork and collaborative environment	2.948	Neutral	1.3926
Teachers' knowledge and teaching methods are up to date in each course.	2.497	Disagree	1.2746
The teacher uses aids in teaching such as (projector, computer, pictures, models, any other ways)	3.919	Agree	1.1443
The teacher uses examples of real life applications to support theoretical foundations.	3.156	Neutral	1.4838
Learning material is published in digital as well as printed form	3.665	Agree	1.2753
There is a suitable number of available references for study material to follow the courses	2.957	Neutral	1.4749
The students are divided into groups to perform projects and assignments with the ability to organize and communicate between members of the Group	3.454	Agree	1.3489

Information is displayed in orderly and clearly form	2.91	Neutral	1.4628
There's a chance to communicate with the teacher out of the class room	2.769	Neutral	1.3805
The teacher uses social media to communicate with students	2.676	Neutral	1.3163
The teacher uses a number of methods such as (tests, seminars, midterm exam) in assess student performance	3.977	Agree	1.1316
Exams questions are clear and cover the content of the course.	2.725	Neutral	1.3263
There is diversity of questions (subjective, objective, analysis) in exams and tests	2.928	Neutral	1.399
There is practical semester (in a real working environment) in the curriculum	2.436	Disagree	1.0674
The office hours are adequate	2.431	Disagree	1.3197

C. Academic advising

Academic Advisor offers a suitable number office hours to meet students	2.436	Disagree	1.2454
I can suitably communicate with the academic advisor	2.561	Disagree	1.3135
Academic Advisor suitably helps me in determining my goals	2.472	Disagree	1.3316

D. Student

The study in this college was my desire and decision.	3.431	Agree	1.4415
I can communicate and collaborate with colleagues	4.003	Agree	1.2752
there are groups and teams where i can use my knowledge and exchange/interact with others	3.454	Agree	1.4119
The relationship with my parents is stable	4.558	Strongly agree	0.983
I can communicate with my parents and reflect problems that I am suffering from	4.017	Agree	1.3491
I can pay tuition fees without problems	3.234	Neutral	1.4784
University allows time for diverse activities (political, sports, cultural, …)	2.448	Disagree	1.2081
The University offers suitable places and infrastructure to practice these activities	2.494	Disagree	1.1299
I understand the rules and regulations of the academic system	3.029	Neutral	1.48
I can communicate with persons who are responsible for student affairs	2.483	Disagree	1.279

i am able to work along with the studies and get practical experience	2.65	Neutral	1.2748

E. HEI

Curriculum design and approval processes involve industrial and business experts and students.	2.092	Disagree	1.1882
The curriculum has a suitable sequence for the students.	2.616	Neutral	1.2897
The University provides vocational training opportunities as well as postgraduate studies.	2.668	Neutral	1.4554
University has international research projects and integrates results in teaching.	2.402	Disagree	1.4558
University has research projects with various sectors in country (industry, health…) and integrates results in teaching.	2.26	Disagree	1.3753
The University provides financial support to poor students	2.46	Disagree	1.2785
University provides incentives for superior students	2.934	Neutral	1.43
University has office dedicated to student affairs	3.289	Neutral	1.4417
University has office dedicated to examination	2.882	Neutral	1.4726
When I have a problem I know which council is responsible for solving it	2.561	Disagree	1.2617

F. ICT

Enough Computers available	3.101	Neutral	1.3636
University provides technical support when using IT	2.711	Neutral	1.3885
University provides fast internet	2.734	Neutral	1.2033
University offers learning management system (LMS)	2.179	Disagree	1.1483
I am using the LMS	2.275	Disagree	1.1904
There is IT supported collaborative work environment	2.309	Disagree	1.1671
The use of personal data is securely treated.	2.523	Disagree	1.4247
There is Wi-Fi access all over the campus	2.231	Disagree	0.7253

Discussion

With regard to the first factor related to the environment, most of the students' answers emphasizes there is a problem with the infrastructure in terms of roads, buildings, maintenance and the need to take disabled students/wheel chair drivers into consideration. In addition, there is a problem with the means of transportation.

For the internal environment of the classroom, the results showed that there are some negative points, which is an effect of external disturbance on the sound inside the room and the seats are uncomfortable.

Study environment suffer from some problems as the lack of equipment in the laboratories, and there are not enough modern references in the field of specialization, and lack of environmental support for the group work in the library.

With regard to services within the university from the opinions of the students, we conclude that the health services are not suitable and the services related to the food are not acceptable and the prices of meals are not suitable.

The second factor is related to teaching; there are some points that negatively affect the learning process such as the lecture schedule where it needs to be improved so that it contains time for practicing other activities such as prayer, eating and other activities. Some of students have problem with language used in teaching.

The dimensions of the teaching methods, some of them have a positive impact on the learning process such as the use of aids in the educational process in addition the availability of printed and digital materials and also the use of different assessment methods. Other has negative impact and need an improvement as the support of collaborative work, the channel of communication and the examinations setting.

The third factor related to academic advising, Academic guidance needs to be well-developed, with responses from half of the students confirming that the specified hours are inappropriate and there is a problem with the communication and help provided by advisor.

The fourth factor about Student, some students suffer from some problems in paying tuition fees and inability to work with the study. Students see that the university does not provide financial support to poor students.In addition, some students suffer from a lack of understanding of the rules and regulations and the difficulty of reaching

out with the student affairs. The practice of activities shows students the lack of time and place to practice other activities.

The five factor related to HEI, There are some negative points where the students' views showed that the sequence of the curriculum is inappropriate and there is no practical semester that allows the student to work in real work environment, in addition to that most students see that thestudents and experts of industry and business are not included in the design of the curriculum.

With regard to research projects within or outside the country, the views varied between agree and disagree of the availability of research projects. Training opportunities students see the there is no training opportunities.

The last factor about the ICT support, the negative points are there is no technical support when using IT, slow internet, there is no WI-FI access and unavailability of LMS all this points affect negatively on the learning process and need an improvement.

All the above negative points need an attention from decision maker in university to improve the learning process.

Appendix-C Questionnaire in Arabic

إستبيان عن العوامل المؤثرة على العملية التعليمية في مؤسسات التعليم العالي

السنة	**النوع**
الأولى ☐ الثانية ☐ الثالثة ☐ الرابعة ☐	ذكر ☐ أنثى ☐
الكلية:	**القسم:**

#	***العبارة***	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
	البيئة					
1	الطرق داخل الجامعة ممهدة					
2	المباني متجاورة ولا اجد مشكلة في التنقل و الحركة بينها					
3	وسائل المواصلات متوفرة و لا اجد صعوبة في الوصول للجامعة					
4	لمباني و الطرق تصان جيداً					
5	القاعات مكيفة و التهوية جيدة					
6	لا يوجد تأثير للضوضاء الخارجية داخل القاعات					
7	الإضاءة داخل القاعات جيدة					
8	الصوت داخل القاعات يسمع بوضوح					
9	عدد الطلاب داخل القاعة مناسب					
10	المقاعد داخل القاعات مريحة					
11	المعامل مجهزة و المعدات متوفرة					
12	المكتبة مجهزة ومريحة					
13	تحتوي المكتبة على المراجع الحديثة في مجال التخصص					
14	توجد غرف داخل المكتبة مخصصة لعمل المجموعات (وبالتالي لن يتم ازعاج االآخرين)					
15	توفر الجامعة خدمة صحية و اسعافية ملائمة					

16	تقدم الكافتريات خدمات جيدة					
17	أسعار الوجبات مناسبة					
18	يستطيع ذوي الاحتياجات الخاصة الوصول لجميع المباني (قاعات-مكتبات-معامل) و الحصول على جميع الخدمات					
	التدريس					
19	جدول المحاضرات تم اعداده بصورة جيدة					
20	يوجد زمن في الجدول للقيام بنشاطات اخري (الأكل، الصلاة، ..)					
21	لأ اجد مشكلة في اللغة (الانجليزية) المستخدمة في التدريس					
22	طريقة الاستاذ في التدريس تناسب طريقتي في التعلم					
23	يتم التدريس بالطريقة التقليدية (اعطاء المعلومة بواسطة الاستاذ فقط)					
24	يوفر الاستاذ بيئة تعليمية تسمح بالتعبير و النقاش و تبادل الاراء					
25	يشجع الاستاذ العمل الجماعي و يوفر بيئة تعاونية					
26	معرفة الاساتذة و طرق التدريس مواكبة في جميع المواد					
27	يستخدم الاستاذ وسائل مساعدة في التدريس (كمبيوتر، بروجكتر، صور، مجسمات ، اي وسائل اخرى)					
28	يستخدم الاستاذ أمثلة واقعية لدعم الأسس النظرية					
29	تتوفر المادة التعليمية في شكل مطبوع و الكتروني					
30	يوجد عدد مناسب من المراجع المتاحة لمتابعة دراسة المواد					
31	يتم تقسيم الطلبة لمجموعات لاداء المشاريع و الواجبات ويوجد طريقة للتنظيم و التواصل ما بين افراد المجموعة					
32	يتم عرض المعلومات بصورة و اضحة و منظمة					
33	هنالك فرصة للتواصل مع الأستاذ خارج القاعة					
34	يستخدم الأستاذ و سائل التواصل الإجتماعي للتواصل مع الطلاب					
35	يستخدام الاستاذ عدد من الطرق (اختبارات، سمنارات، امتحانات منتصف فصل) في تقييم الأداء في المادة					
36	اسئلة الامتحانات واضحة و شاملة لمحتوى المقررات					
37	يوجد تنوع في أسئلة الامتحانات و الاختبارات (اسئلة في الموضوع و تحليله)					

38	يوجد فصل دراسي عملي(في بيئة عمل حقيقية خارج الجامعة) في المنهج					
39	الساعات المكتبية مناسبة					
	الإرشاد الأكاديمي					
40	يحدد المرشد الأكاديمي عدد مناسب من الساعات المكتبية لمقابلة الطلاب					
41	استطيع التواصل بطريقة مناسبة مع المرشد الأكاديمي					
42	المرشد الأكاديمي يساعدني بطريقة مناسبة في تحديد اهدافي					
	الطالب					
43	الدراسة في هذه الكلية رغبتي و قراري					
44	استطيع التواصل و التعاون مع زملائي					
45	يوجد فرق و مجموعات استطيع التفاعل و النقاش و تبادل المعرفة مع الاخرين					
46	العلاقة بين الأبوين مستقرة					
47	استطيع التواصل مع الوالدين و اعكس المشاكل التي اعاني منها					
48	استطيع سداد الرسوم الدراسية دون مشاكل					
49	تتيح الجامعة زمن لممارسة النشاطات (سياسية، رياضية، ثقافية، ...)					
50	توفر الجامعة اماكن مناسبة لممارسة هذه النشاطات					
51	اعرف و افهم اللوائح المتعلقة بالنظام الأكاديمي					
52	استطيع التواصل مع الجهات المتعلقة بشؤون الطلاب					
53	انا قادر على العمل جنبا الى جنب مع الدراسة و استطيع الحصول على الخبرة العملية					
	مؤسسة التعليم العالي					
54	تصميم المناهج و الموافقة عليها يتم بإشراك خبراء الصناعة و الاعمال و الطلاب					
55	المنهج متسلسل بطريقة مناسبة للطلاب					
56	توفر الجامعة فرص للتدريب المهني و الدراسات العليا					
57	تمتلك الجامعة مشاريع بحثية عالمية و تدمج نتائج هذه المشاريع في التدريس					
58	لدى الجامعة مشاريع بحثية مع مختلف القطاعات في الدولة					

	(الصناعة، الصحة، ....) و تدمج نتائج هذه المشاريع في التدريس					
59	توفر الجامعة دعم مالي للطلاب المتعسرين					
60	تقدم الجامعة حوافز للطلاب المتفوقين					
61	يوجد بالجامعة مكتب متعلق بشؤون الطلاب					
62	يوجد مكتب متعلق بالإمتحانات					
63	عندما تواجهني مشكلة اعرف الجهة المسؤولة عن حلها					
	تكنولوجيا المعلومات و الاتصالات					
64	أجهزة الكمبيوتر متوفرة و متاحة					
65	توفر الجامعة دعم تقني لاستخدام تكنولوجيا المعلومات					
66	توفر الجامعة شبكة انترنت سريع					
67	يوجد بالجامعة نظام ادارة تعلم الكتروني (LMS)					
68	استطيع استخدام نظام ادارة التعلم					
69	يوجد دعم تقني لبيئة العمل الجماعي					
70	يتم التعامل مع البيانات الشخصية بشكل آمن					
71	تتوفر خدمة ال Wi-Fi في جميع انحاء الجامعة					

Publication one

Investigating Factors Affecting the Learning Process in Higher Education - A Case Study

Wafa Ali Mohammed
Computer Science Department, PhD program
Sudan University of Science and Technology
Khartoum, Sudan
wafaali@sustech.edu

Abstract — ***Globalization, Information and Communication Technology (ICT) advances, and the expansion in Higher Education (HE) pose many challenges on HE Institutions (HEIs). Chief among them is gaining a competitive advantage necessary for sustainability, achieving goals and providing a service that satisfies customers (students). Learning is the most important service provided by HEIs, and it involves two other stakeholders in addition to the institutions, namely, students and educators. The aim of this study is to identify the main factors that affect the learning process at HEIs and to empirically investigate their effectiveness. Content analysis techniques and exploratory interviews were used in identifying main factors and their associated variables. Data mining was utilized for analysis to discover patterns and extract valuable information. The case study involves students from the College of Computer Science and Information Technology (CSIT) at Sudan University of Science and Technology (SUST). A survey instrument was used to gauge stakeholders' perception and expectation of the effect of the predetermined factors. The results showed that many factors have an adverse impact on the learning process and require attention, including transportation, ICT support, and services such as cafeteria and health services, etc. This research contends that identifying and highlighting the effect of factors from the perspective of the main "customers" would yield better students' satisfaction, reduced dropout rates, enhanced HEIs competitiveness and eventually improved learning process.***

Keywords — Learning process, higher education, ICT, decision-making, data mining, clustering.

I. INTRODUCTION

Education is defined as a learning process for the individual to attain knowledge. It is an important process for individuals and society since it contributes to the progress and development of societies [1], [2]. No society can advance and become more civilized without education. The culture and values of any society depend on the education of individuals as cultural values are passed on through educational institutions. Generally, the formal educational process is composed of the following stages, kindergarten, primary, middle, secondary and the HE stage. HEIs deal with post-secondary level of education, and this includes undergraduate and postgraduate colleges, universities and centers.

Numerous research [3], [4] has shown that the core functions of HEIs are: education, research and contributing to the society. In addition, they play a vital role of supplier to the labor market, as they provide highly qualified human resources. Learning is the most important process in HEIs, and it involves three main stakeholders, students, educators, and institutions.

HEIs currently face a variety of challenges, as the spread of ICT and globalization has unlatched international competition in addition to the local one. Moreover, the surge in the number of students without financial support from governments, and the expansion in HEIs both locally and globally. Confronting these challenges requires strategic and tactical planning by HEIs as they must maintain their existing students and stay competitive in recruitment by providing a service that satisfies their customers (students). HEIs must also respond to and stay in-line with the continuous changes in the labor market. Focusing on learning and how this service can be provided in a modern, efficient, and satisfactory to students will make HEIs more competitive.

Innovations in ICT have enabled HEIs to better respond to their growing challenges and achieve their objectives. Currently, almost all HEIs utilize ICT in different capacities to support students' operations, HE management, administrative operations, and planning. Many research articles [5], [6] were published in recent years on improving the learning process in HEIs. Similarly, there are several studies on the use of data analysis including data mining techniques in supporting decision-making at HEIs [7], [8], [9], [10]. This paper combines these two research areas by

utilizing data mining techniques to enhance the learning process.

To improve the learning process, first, an identification and examination of the factors affecting its success must be carried out. Followed by an investigation of stakeholders' perception. From the previous discussion, the following research questions arise: what are the main factors that affect the learning process? How can HEIs improve the learning process? This study's objective is to identify and evaluate the factors that influence the learning process from the perspective of students. The study focuses on students because they are the main stakeholders and their satisfaction is the target of HEIs**.** For this purpose, a survey instrument was administered to CSIT students at SUST in a classroom setting.

In order to deliver a meaningful analysis, data mining clustering techniques were applied to provide further knowledge beyond the data [11], [12]. The aim is to discover structure inside unstructured data, extract meaning from noisy data, discover patterns in apparently random data, and use all this information to better understand trends, patterns, correlations, and ultimately predict students' behavior.

The paper is organized in five sections beginning with this introduction. The following section presents background information about learning, the learning process, and data mining. The next section outlines the research methodology utilized to identify factors affecting the learning process, creating a survey instrument, and the analysis approach. The fourth section presents the survey results and discusses their significance. While the last section provides a summary of the research and proposes future direction.

II. BACKGROUND

Learning is an essential activity and is even more vital for youth, because through its institutions they prepare for their future roles in the development of their societies and acquisition of success in life [13], [14]. There are many definitions for learning, the Oxford Companion to Philosophy defines it as "the acquisition of a form of knowledge or ability through the use of experience". Another one is that "learning is an activity with purpose of making change in behavior, attitude, habits, knowledge, skill, etc." [15].

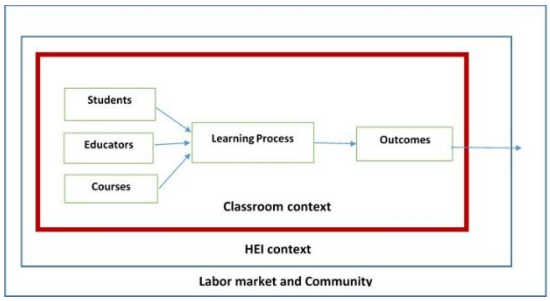


Fig. 29. Learning process

Learning is done through providing and improving the learning process. The learning process is comprised of the interaction between institution, educator, student, environment and courses. The outputs from the learning process are the graduate students and their quality [16], [17]. Fig. 1 gives a general view of the learning process and the interaction between its elements. The practice of learning is done in a classroom context. This could be a physical classroom in face-to-face traditional education or virtual classroom in e-learning. The classroom setting affects the ongoing learning process. Classrooms are generally inside the HEI context, where all regulations, policies, environment and facilities provided by the institution affect the learning process. There is a strong bidirectional relation between the labor market and the HEI context, as labor market and community requirements affect the design of curriculum to produce graduates who in turn contribute to community and business development. The community is affected by the final output, because the development of community in all aspects such as, economic, political, social and cultural depends on the outputs of the learning process.

From the preceding information and discussion, we conclude that the learning process is complex, interconnected and involves many factors and stakeholders. Identifying these factors and how they affect the learning process is an important issue for any improvement. Similarly, stakeholders play a role in the success or failure of the process, with students being the most important one. For two reasons, first they are at the heart of the process being the main customers for HEIs. Second, their success is the target and outcome of the whole process as graduates. Therefore, any effort to improve the learning process must take into account the view point of students. This includes investigating their attitude, expectation and perception toward various specific activities and characteristic features of the process, along with socioeconomic factors. The next section presents and explains the factors affecting the learning process in more detail.

Numerous recent research articles have proposed utilizing ICT tools and techniques to assist in HEIs management and planning. For example, using ICT to support management process [18], to help with decision-making [19],

staff development and curriculum design [20], and for general administration (payroll, administration of students' data, etc.) [21]. in this study, data clustering technique is utilized to analyze the case study survey results. Clustering is an unsupervised learning data mining method that has become an essential component to various organizations due to its significance in decision-making support.Data mining is the process of analyzing data from different perspectives and summarizing it into important information so as to identify hidden patterns from a large data set [22].

Clustering is the process of partitioning a set of data into a set of meaningful groups, called clusters. Clustering algorithms can be used to find natural groupings when there are many cases and no obvious groupings. Clustering relevant data in several clusters makes it easy to analyze. In this study we explore and analyze the factors that affect the learning process from students view. Thus, partitioning students in groups based on their characteristics and how they asses these factors, simplifies data analysis and the understanding the problem.

The widely used clustering algorithm, K-means, is the one selected for data analysis. It is considered a hard-clustering algorithm since each object is assigned to one cluster. The k-means algorithm is simple and easy to understand [23].

III. METHODOLOGY

This section describes the methodology used to realize the research objectives by answering the research questions. It is organized into two phases: in the first phase content analysis and analytical research procedures are used while performing an extensive review of literature to identify the factors that affect the learning process. In the second phase a questionnaire is designed to gauge the effect of the previously identified factors on the process.

A. Factors Affecting the Learning Process

In the first phase, we review the literature to explore, identify and analyze in detail the learning process, factors affecting it, and stakeholders involved. Consequently, and as outlined in the previous section, the process of learning is influenced by many factors related to student, family, class, colleagues, educators, courses, teaching methods, technology, HEI policy, regulations, aids and the general environment. These factors can be grouped into six mainfactor groups or dimensions, namely: environment, student, teaching, academic advising,HEI, and ICT.

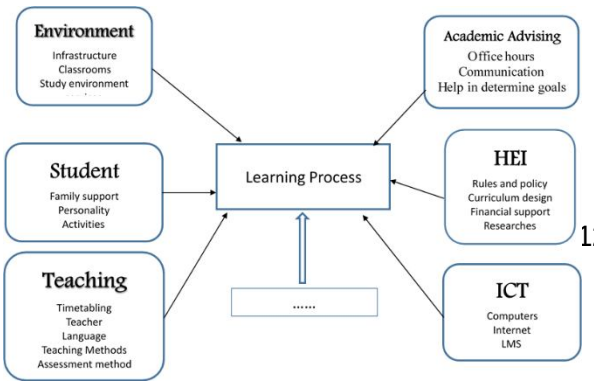


Fig. 2. Factors affecting learning process

Fig. 2 presents these six dimensions with arrays of their comprising factors or characteristics. These arrays are dynamic and can expand or reduce in size over time.

1) ***Environment:***Thelearning environment refers to the space allocated for classrooms, science labs, open spaces and offices. It also refers to diverse physical locations, contexts and culture in which students learn [24]. From this definition we can see that environment is about different physical aspects that affect the learning process inside the classroom, e.g., light, sound, air condition, seats, etc., and outside the classroom. The physical classroom environment is one area that makes an obvious impact on students' success because most of their time is spent there. There is a clear relationship between the physical structure of classroom and the psychological factors as behavior and interaction in the classroom [25].

The HEI environment outside the classroom, such as roads, garden, parking, adjacent buildings, easy movement, etc., all these things affect the behavior of students and their achievement and their motivation towards study.

2) ***Student:*** The student plays an important role in the success of the learning process. There are many student characteristics that influence the process. First, is motivation or desire, and this the driving force behind student learning and success [26]. Second, is family environment and support. This is characteristic is reflected in students' relation to their parents, their communication with them, and their family income. Students who live in stable family environment and find good support, do well in their studies. Third, the relationship and communication with colleagues.

3) ***Teaching:***This dimension contains five important factors, the first and foremost factor is the educator. The role of an educator at HEIs is more like a facilitator than teacher, as she/he leads students through the learning process. Educators play an extremely important role in guiding students to achieve their academic goal [25]. One of the basic elements that influence the learning process is educators having solid and up-to-date knowledge, and encouraging interaction with students in and out of the classroom. There are technical and personal requirements in the teaching factor, such as: appropriate verbal instructions, that can be interpreted as the method of teaching, feedback to students about their progress, good relationship with the student, and attitude towards the subject taught [26].

The second factors in the teaching dimension is teaching aids, and they are those materials that help educators explain better and help students

understand. Teaching aids play an important role in the learning process, as they motivate students, clarify the information, and make the classroom alive and active. They are many types of teaching aids such as visual, audio, and audiovisual aids [27].

The third factor is courses, sequence of courses, contents, and curriculum design and if it relates to the labor market. All of these aspects influence the learning process. Courses content should meet students' need and their objective must be clear [17]. Related to contents is how the information is displayed or presented to students, the educator must take into account the different learning styles in the classroom [28]. Learning styles refers to the preferred way an individual processes information and also describes a person's typical mode of thinking, remembering or problem solving. The basic learning styles are visual, auditory and Kinesthetic [29].

Courses assessment methods enable educators to ensure that students learn what they need to meet courses learning objectives. They include strategies, techniques, tools and instruments for collecting information to determine the extent to which students demonstrate the desired learning outcomes. Using multiple assessment methods leads to a complete view about students' achievement.

The fourth factor in this dimension is teaching method, and this is the method of subject delivery to students. Teaching methods must correspond to students' characteristics. There are different kinds of teaching methods, including lecture, discussion, question and answer and group work. Choosing the appropriate method leads to increased learning achievement [30]. The last factor in the teaching dimension is the lectures timetable. Timetable setting is an important issue, for example, if the timetable is compressed and does not contain any gaps between lectures, it negatively affects students' comprehension. Students need time between lectures to perform other activities such as eating, drinking, praying, etc. They also need to relax their mind so receiving new information could be easy.

4) ***Academic Advice:*** Advising is a process in which advisor and advisee enter a dynamic relationship respectful of the student's concerns. Ideally, the advisor serves as educator and guide in an interactive partnership aimed at enhancing the student's self-awareness and fulfillment [31]. Advising process can help students to identify personal strengths and interests related to their educational and career goals [32]. Academic advising is important because it is a central key to any student's success in college. Advisers monitor students' academic progress, give students the ability to navigate the HEI policies and procedures better, and assist students in their problems. Continuous interactions between students and their academic advisor isimportant and helps in increasing student retention [33], and this plays a vital role in improving the learning process.

5) ***HEI:*** The HEI itself plays a key role in the learning process by designing curricula related to the labor market and providing the necessary environment for students and educators. HEI must also establish clear and understandable policies and regulations for students, offer incentives to superior students, while providing financial assistance to poor students. Moreover, HEI develop research projects and reflect their results on the whole process. All of these characteristic factors directly affect the quality and efficiency of the learning process.

6) ***ICT:*** It is obvious that ICT utilization and its necessary support have strong influence on all stakeholders of the learning process. They do so in a number of ways, such as, facilitating educators' access to information, the preparation of educational materials and the presentation of information to students. ICT provides the possibility of sharing information between educators and students, and students among themselves. It also enables opportunities for discussion and exchangeof ideas through the use of academic sites, social media and collaboration work techniques [34]. Furthermore, the availability of adequate numbers of computers,high speed internet, Wi-Fi, secure data processing and support are extremely important for students and facilitates their learning.

B. Survey Design

In the second phase of the research, a survey was designed to capture students' feedback on the identified factors influencing the learning process. The questionnaire consists of two sections, the first is about personal data to illustrate the characteristics of the sample of the study, including college, department, academic year and gender. The second section is divided into six sub-section corresponding to the dimensions outlined earlier. These sections consist of 71 questions with each dimension covering its related factors.

A five-point Likert scale was used to measure student responses to the questionnaire sections as follow: 1 - "strongly disagree", 2 - "disagree", 3- "neutral", 4 - "agree", 5 - "strongly agree"

The WEKA data mining tool was used for analyzing the survey results. WEKA stands for Waikato Environment for Knowledge Analysis, and it is a data mining or machine learning tool [35] (Bouckaert, Frank et al. 2010) (Bouckaert, Frank et al. 2010). It is an open source tool that supports different standard data mining tasks such as data preprocessing, classification, clustering, regression, visualization and attribute selection.

WEKA is easy to use and has many built-in features that require no programming [36].

IV. RESULTS AND DISCUSSION

The participants in this study are students from the CSIT College at SUST. CSIT has four departments, namely, Computer Sciences (CS), Computer and Information Systems (CIS), Computer Systems and Networks (CSN) and Software Engineering (SE).

TABLE XXII. GENERAL CHARACHTERISTICS

Department	Students	Male	Female
CS	83	36	47
CIS	72	35	37
CSN	79	28	51
SE	113	54	59
Total	347	153	194

As Table I shows, the total number of students who responded to the questionnaire is 347, of those 153 are male (44%) and 194 are female (56%).

The K-means clustering algorithm utilized for data analysis allows users to select the number of clusters. In this study, after exhaustive testing, we decided to use three clusters (k=3) since it gave clear and understandable results. This implies that all analysis results will be divided in three groups based on respondents' feedback. These groups will be called Cluster0, Cluster1 and Cluster2.

1) *Environment*:

TABLE XXIII . ENVIRONMENT DIMENSION

Factor		Cluster0	Cluster1	Cluster2
1	HEI infrastructure	2.3	3.3	2.8
2	Transportation	1.6	2.5	1.8
3	Classroom environment	2.2	4.0	3.1
4	Study environment	2.0	3.5	2.9
5	HEI services	1.8	2.9	2.4

Table II show the results of the five Environment factors and the Mean for the three clusters, i.e., cluster0, cluster1 and cluster2. Each cluster represents a group of students and the Mean is the output from the K-means algorithm. The Mean here represents the weighted average. Since we are using Likert quintet, the range is 4, and the number of options equals 5, then the weighted average becomes 0.8. Therefore, each option Mean falls in the following range: 1-1.79 "strongly disagree", 1.80-2.59 "disagree", 2.60-3.39 "neutral", 3.40-4.19 "agree", and 4.20-5 "strongly agree". It is clear from Table II that "Transportation" is a problem that faces the majority of students. All students have to commute to and from the SUST campus and the lack of transportation affects their morning lectures attendance. The results of "HEI services" too show that most students are not satisfied with services such as healthcare and cafeteria services.

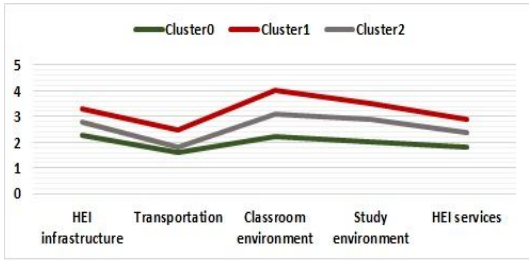


Fig. 3. Environment Dimension

Fig. 3 is a line graph of the Environment dimension results, where the y-axis represents the Mean. From the figure we can clearly see that students grouped in Cluster0 are not satisfied with the environment factor in general. Cluster1 students are mostly neutral butthey are satisfied with the classroom environment. Cluster2 students are dissatisfied with both transportation and HEI services while they are neutral with other factors.

2) *Student:*

TABLE III. STUDENT DIMENSION

Factor	Cluster0	Cluster1	Cluster2
Family support	3.7	4.4	4.1
Personality	2.6	3.8	3.4
Extracurricular	1.7	3.4	1.9

Table III presents the results of the Student dimension. "Family support" explores the student's relation with his/her parents, the ability to communicate with them, and also if the student can pay tuition fees. From Table III results, the general pattern of answers is "agree". This means that most students have good financial and moral support from their families.

Fig. 4. Student Dimension

Fig. 4 is a bar-chart representation of the Student dimension and the y-axis represent the Mean. The second

factor, i.e. personality, exposes the students' desire to study, communicate with colleagues, and communicate with people responsible of his/her affaires. From the results in Fig. 4, we can see that some of students have problems with this factor. The third factor deals with extracurricular activities such as political, sports, cultural, etc. It addresses the availability of time and infrastructure to practice such activities on the SUST campus. The results in the table and graph clearly demonstrate that SUST needs to put more emphasis on supporting these activities as they positively effect students and the learning.

3) Factor3: Teaching

TABLE IV. XXIV TEACHING DIMENSION

Factor	Cluster0	Cluster1	Cluster2
Lectures Timetable	1.4	2.8	1.9
Language barrier	2.3	3.4	2.8
Teaching method	2.3	3.4	2.8
Teaching support	2.8	4.1	3.6
Communication	2	3.5	2.8
Collaborative work	2	3.8	3.2

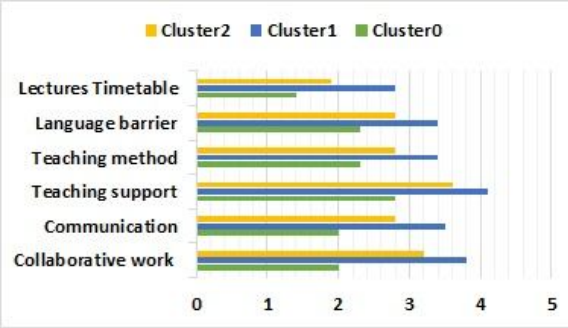


Fig. 5 Teaching Dimention

Result of the analysis of the Teaching dimension and its factors is presented in Table IV and Fig. 5. Generally, most responding students ranked all factors low except the students in Cluster1, they ranked "Teaching support" and "Collaborative work" high. These results reflect the strong dissatisfaction of students with the teaching dimension, and have negative impact on the learning process. Consequently, there is an urgent need for improvements as the teaching factors play vital role in learning process and have direct relation with students' satisfaction.

4) Academic Advice:

TABLE V. ACADEMIC ADVICE

Factor	Cluster0	Cluster1	Cluster2
Office hours	1.2	3.7	2.2
Communication	3.5	3.4	1.7
Help to determine goals	3.2	3.6	1.5

With regard to this dimension, students were asked about the office hours for meeting their advisor, the communication with the advisor, and if they get adequate assistance from her/him to achieve their goals. From the result presented in Table V, there are varying views between dissatisfied and satisfied, and this is due to the evaluation of different advisors. This indicates that there is a difference in performance between academic supervisors.

5) HEI:

TABLE XXV. HEI DIMENSION

Factor		Cluster0	Cluster1	Cluster2
1	Curriculum design	2.9	2.4	2.0
2	Additional education opportunities	3.9	1.9	1.9
3	Financial support	3.4	3.0	1.9
4	Students' affairs	3.7	3.3	1.9

HEI dimension has four factors as shown in Table VI. The "Curriculum design" factor addresses the process of course design and the sequencing of courses. From the results in Table VI, the majority of students are not satisfied with this factor since the Mean value is less than 3 in all clusters, meaning that it negatively affects the learning process. The second factor in the table explores if the HEI provides additional education opportunities, such as professional or summer training, internship support, international research projects, etc. This factor too has a negative effect on learning process because students in Cluster1 and Cluster2 are not satisfied. The third factor in this dimension, deals with the financial support for students. Here responses vary, as some students see that SUST provides enough support while others disagree. The last factor addresses students' affairs, i.e., do students know where to go if they have a problem or issue, and to whom do they talk. Again, in this factor there is no agreement. From the discussed results we conclude that there are some factors in the HEI dimension that need improvement.

6) ICT:

TABLE XXVI. ICT DIMENSION

Factor	Cluster0	Cluster1	Cluster2
Technical support	4.0	4	2.1
Network and internet Availability	3.6	2.6	2.2
Availability of LMS	3.6	1.8	2.1

From Table VII, most students are satisfied with the first factor, i.e., availability and accessibility of computers and technical support. At the same time, most are not satisfied with the second factor, i.e., internet, network, and Wi-Fi

availability, the Mean value is less than 4 in all clusters. This implies that there is a need to provide better internet connectivity with high speed and Wi-Fi connection in order to improve the learning process. The last factor is about the availability of a Learning Management System (LMS), and the majority of students don't know about LMS and have not used it. ICT is important factor, and there is relation between ICT and other factors. Supporting and encourage the using of ICT will lead to improve the learning process.

V. CONCLUSION

This paper has investigated the factors that affect the learning process at HEIs. It has identified essential factors and empirically examined the core stakeholders' (students) feedback. A survey instrument was used to collect data from the target group, i.e., CSIT students at SUST. Analysis of the survey results has shown that there are various factors that adversely affect the learning process at SUST. For example, transportation, HEI services, including cafeteria and healthcare services, lectures timetable, academic advice, curriculum sequences and ICT support. These factors require further in-depth study, focus and analysis by SUST administrators to explore ways to improve performance and thus enhance the learning process. Furthermore, the research has shown that ICT support is a vital factor, not only because it impacts all stakeholders, but also for its influence on other factors as well.

This study has laid the foundation for improving the learning process, and achieved the outlined research objective, while demonstrating that the use of data analysis tools and techniques is a key component in supporting HEIs administrators and planners in their decision-making. It has also provided the base for further research and extending to other data analysis techniques, case studies, and other related contexts.

REFERENCES

[1] The World Bank, "Higher Education in Developing Countries: Peril and Promise," Washington, DC, 2000, p. 138.

[2] F. Teal, "Higher Education and Economic Development in Africa: A Review of Channels and Interactions," Journal of African Economies 20, no. Suppl. 3, 2011, 50-79.

[3] JICA (2004). "Approaches for Systematic Planning of Development Projects/higher education."

[4] Danil, SIR John. "Higher Education: Past, Present, and Future--A View from UNESCO." *Higher Education in Europe* 28, no. 1 (2003): 21-26.

[5] Crosling, Glenda, Margaret Heagney, and Liz Thomas. "Improving student retention in higher education: Improving teaching and learning." *Australian Universities' Review, The* 51, no. 2 (2009): 9.

[6] Crosling, Glenda, Margaret Heagney, and Liz Thomas. "Improving student retention in higher education: Improving teaching and learning." *Australian Universities' Review, The* 51, no. 2 (2009): 9.

[7] Zhang, Ying, Samia Oussena, Tony Clark, and Kim Hyensook. "Using data mining to improve student retention in HE: a case study." (2010).

[8] Agaoglu, Mustafa. "Predicting instructor performance using data mining techniques in higher education." *IEEE Access* 4 (2016): 2379-2387.

[9] Raju, Dheeraj, and Randall Schumacker. "Exploring student characteristics of retention that lead to graduation in higher education using data mining models." *Journal of College Student Retention: Research, Theory & Practice* 16, no. 4 (2015): 563-591.

[10] Asif, Raheela, Agathe Merceron, Syed Abbas Ali, and Najmi Ghani Haider. "Analyzing undergraduate students' performance using educational data mining." *Computers & Education* 113 (2017): 177-194.

[11] Berkhin, Pavel. "A survey of clustering data mining techniques." In *Grouping multidimensional data*, pp. 25-71. Springer, Berlin, Heidelberg, 2006.

[12] Gulati, Hina, and P. K. Singh. "Clustering techniques in data mining: A comparison." In *2015 2nd international conference on computing for sustainable global development (INDIACom)*, pp. 410-415. IEEE, 2015.

[13] Siming, Luo. "Factors Leading to Students' Satisfaction in the Higher Learning Institutions." *Journal of Education and Practice* 6, no. 31 (2015):114-118..

[14] Khurshid, Fauzia. "Factors affecting higher education students' success." *Asia Pacific Journal of Education, Arts and Sciences* 1, no. 5 (2014): 40-47.

[15] Lachman, Sheldon J. "Learning is a process: Toward an improved definition of learning." *The Journal of psychology* 131, no. 5 (1997): 477-480.

[16] Betoret, Fernando Doménech, and Adela Descals Tomás. "Evaluation of the university teaching/learning process for the improvement of quality in higher education." *Assessment & evaluation in higher education* 28, no. 2 (2003): 165-178..

[17] Pavione, Caroline Stéffani Santos Nério, Bruna Camargos Avelino, and José Roberto de Souza Francisco. "Factors that Influence the Teaching-Learning Process from the Perspective of Accountancy Students: Analysis at a Higher Education Institution in Minas Gerais." *Revista de Educação e Pesquisa em Contabilidade* 10, no. 2 (2016).

[18] Ujunju, Michael Okumu, G. Wanyembi, and Franklin Wabwoba. "Evaluating the role of information and communication technology (ICT) support towards processes of management in institutions of higher learning." (2012).

[19] Laudon, Kenneth C. *Management information systems: Managing the digital firm*. Pearson Education India, 2007.

[20] Mostert, Markus, and Lynn Quinn. "Using ICTs in teaching and learning: Reflections on professional development of academic staff." *International Journal of Education and Development using ICT* 5, no. 5 (2009): 72-84.

[21] Krishnaveni, R., and J. Meenakumari. "Usage of ICT for Information Administration in Higher education Institutions–A study." *International Journal of environmental science and development* 1, no. 3 (2010): 282-286.

[22] Borkar, Suchita, and K. Rajeswari. "Predicting students academic performance using education data mining." *International Journal of Computer Science and Mobile Computing* 2, no. 7 (2013): 273-279.

[23] Solanki, Harshvardhan. "Comparative study of data mining tools and analysis with unified data mining theory." *International Journal of Computer Applications* 75, no. 16 (2013): 23-28.

[24] Amirul, Nurul Jannah, C. N. Che Ahmad, Asmayati Yahya, M. F. N. L. Abdullah, Mazlini Adnan, and N. Mohamed Noh. "The physical classroom learning environment." In *Proceedings of the International Higher Education Teaching and Learning Conference*, vol. 2, no. 1, pp. 1-9. 2013.

[25] Puteh, Marzita, Che Nidzam Che Ahmad, N. Mohamed Noh, Mazlini Adnan, and Mohd Hairy Ibrahim. "The classroom physical environment and its relation to teaching and learning comfort level." *International Journal of Social Science and Humanity* 5, no. 3 (2015): 237-240.

[26] Aslam, Hassan Danial, A. Younis, A. A. Sheik, M. Z. A. Maher, and Z. A. Abbasi. "Analyzing factors affecting students' satisfaction regarding semester system in universities of Pakistan." *Journal of American Science* 8, no. 10 (2012): 163-170.

[27] Nikky, S. "Characteristics of Good Teaching Aids." *Retrieved December* 4 (2010): 2012.

[28] Abante, Mark Enrick R., Benjie C. Almendral, Jay-ren E. Manansala, and Jovielyn Mañibo. "Learning styles and factors affecting the learning of general engineering students." *International Journal of Academic Research in Progressive Education and Development* 3, no. 1 (2014): 16-27.

[29] Abante, Mark Enrick R., Benjie C. Almendral, Jay-ren E. Manansala, and Jovielyn Mañibo. "Learning styles and factors affecting the learning of general engineering students." *International Journal of Academic Research in Progressive Education and Development* 3, no. 1 (2014): 16-27.

[30] Cook, Sandra. "A chronology of academic advising in America." *The Mentor: An Academic Advising Journal* 1, no. 2 (1999).

[31] Young-Jones, Adena D., Tracie D. Burt, Stephanie Dixon, and Melissa J. Hawthorne. "Academic advising: does it really impact student success?." *Quality Assurance in Education* (2013).

[32] Lau, Linda K. "Institutional factors affecting student retention." *Education-Indianapolis then Chula Vista*- 124, no. 1 (2003): 126-136.

[33] Munawaroh, N. "The influence of teaching methods and learning environment to the student's learning achievement of craft and entrepreneurship subjects at vocational high school." *International Journal of Environmental & Science Education* 12, no. 4 (2017): 665-678.

[34] Kumar, DTT Vijaya, J. Senthil, and Ramesh Palanisami. "*Optimizing Resources in Teaching and Learning using Cloud Technology*."

[35] Han, Jiawei, Jian Pei, and Micheline Kamber. *Data mining: concepts and techniques*. Elsevier, 2011.

[36] Bouckaert, Remco R., Eibe Frank, Mark A. Hall, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. "WEKAâˆ'Experiences with a Java Open-Source Project." *Journal of Machine Learning Research* 11, no. Sep (2010): 2533-2541.

Publication two

A Business Intelligence Solution for Improving the Higher Educational Learning Process

Wafa Ali Mohammed[#1], Izzeldin A. Elhassan[#2]
Computer Science Department, PhD program
Sudan University of Science and Technology
Khartoum, Sudan
[1]wafaali@sustech.edu
[2]izzo.elhassan@gmail.com

Abstract — ***Higher Education Institutions (HEIs) today face many challenges such as the local and international expansion and the widespread utilization of Information and Communication Technologies (ICT) that facilitate global competition. To deal with these and with the changes in labor market, HEIs administrators and planners need information to improve their services, satisfy customers (students) and gain competitive advantage. Learning is a process that includes interactions between students, educators and courses under the umbrella of HEIs. In fact, it is the core and most important service provided by them and it is surrounded by many factors. Improving the learning process is multifaceted as there are numerous factors and stakeholders involved. The aim of this study is to present a Business Intelligence (BI) solution and investigate its value in streamlining the learning process at HEIs. The proposed solution avails information and analytical capabilities to HEIs decision makers. A mix of research methods were used to validate the solution, including content analysis to identify and categorize the factors involved in the learning process, along with a case study implementation of the proposed BI solution. Data was collected from different departments at the Sudan University of Science and Technology (SUST). The data analysis clearly supports the research hypothesis that utilizing BI yields reliable information, knowledge-based decisions, and hence improved learning process.***

Keywords — Learning process, higher education, decision-making, business intelligence, data warehouse, dimensional model.

I. INTRODUCTION

Numerous research [1], [2] has shown that the core functions of HEIs are: education, research and contributing to the society. Moreover, HEIs play a vital role of supplier to the labor market, as they provide highly qualified human resources. Learning is the most important process at HEIs, and it involves three main stakeholders, students, educators, and administrators [3]. The learning process is a system that consists of input, processing and output. The inputs are: students, educators, courses, and environment. The processing in the system is the teaching process and the outputs are the graduate students, their quality, and effect on community and labormarket [4], [5].

Advents in ICT have led to more and more data being generated and stored. Then organizations realized the potential value that resides in their data, and subsequently they started to explore ways to utilize this valuable asset. Therefore, high demand arose for ICT methods and tools that can constructively transform data into information and knowledge. Decision-making plays a vital role within organizations and many times they are the reason behind their success or failure. Using systems to support the decision-making processes is significant since they help in data collection, organization, analysis and transformation to information [6].

BI offers emerging solutions for improving data analysis, decision-making and trend investigation. It plays a vital role since it improves the timeliness and quality of information, and enhances communication among departments while coordinating activities, and enabling quick responses to changes whether in financial conditions, customer preferences, or supply chain operations [7]. BI gives organizations a comprehensive and integrated view of their business operation and potential, and this facilitates the decision-making processes. Currently it's difficult to find successful organizations that do not yet use BI [8]. BI is used in different fields such as healthcare, communication, marketing and finally it has entered the education filed.

Clearly, there are many factors that affect and interact with the learning process and its many stakeholders. To improve the learning at HEIs a deep study of the process and its influencers must be carried out. Recent research [4], [5], has shown that these factors can be grouped into SIX main groups or dimensions, namely: *environment*, *student*, *teaching*, *academic advising*, *HEI*, and *ICT*. This leads to this research's question: how

effectively create the data models toward an easy-to-implement scalable BI framework leading to improvements in the educational processes based on these learning process influences?

The objective of this research is to design a BI framework that analyzes the data collected toward new opportunities for managing and improving HE processes, and to empirically investigate its effectiveness. The research contends that adopting BI increases the quality of information, reduces time, and increases the efficiency.

This paper is organized in five sections beginning with this introduction. The following section presents background information about the learning process at HEIs, decision-making, BI, its component and adoption at HEIs. The next section describes the research methodology that is organized in three phases. The first phase, identifies and categorizes the factors involved in the learning process and BI components. While the second phase, explores the development of a BI model that takes advantage of the identified learning process influences and the data collected. The third and final phase, is the data analysis phase which discusses in detail the extraction of knowledge and validation of the model. Finally, the last section provides a summary of the paper and addresses the future direction of the research.

II. BACKGROUND

Globalization and the widespread use of ICT have led to increased worldwide competition between HEIs. Consequently, the need for competitive advantage has made many HEIs tend to use BI technologies in assisting with decision-making and solving problems. Many recent studies were published about BI utilization and applications at HEIs. For example, there are studies about using BI to assist in analytical activities regarding student affairs [9], and for analyzing students' performance [10], and enhancing the teaching-learning process [11]. Additionally, there is research on designing a BI framework for HEIs decision makers [12], building an academic DW [13], [14], and aligning HEIs output with labor markets [15].

The learning process is vital and must always adapt with the labor market changes especially in this global knowledge economy. Improving the process is done by studying all the involved factors and identifying weaknesses and trying to resolve and improve them. As indicated earlier in the introduction, the list of factors influencing the process is not endless and is scalable, as the world around the learning process changes, the list can increase over time [5].

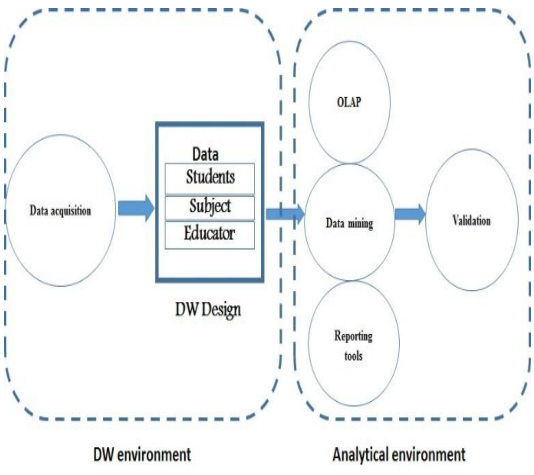
Fig. 1. Proposed BI solution

Fig. 1 illustrates the main components of the proposed BI framework that is comprised of two environments, DW and analytical. The DW environment contains the operational systems that supply the core data, the Extract, Transform and Load (ETL) process, and the DW creation. The analytical environment is where the user creates reports and visual aids, queries the system for specific requests, and makes decisions based on information from the DW.

To gain greater understanding of the components of the learning process and the factors affecting the success of the proposed BI model, the next section explores the creation of a case study implementation of the model at SUST together with the approach for its validation.

III. METHODOLOGY

This section describes the methodology used to realize the research objective by answering the research question. Content analysis, analytical research procedures, and exploratory interviews were used while performing an extensive review of literature to identify the factors that affect the learning process, stakeholder, and data sources required for building the BI solution.



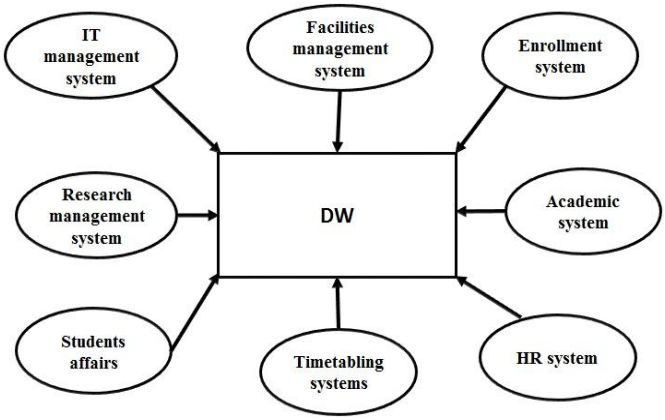


Fig. 2. Operational system

Based on the factors influencing the learning process outlined earlier in the introduction, the data for the proposed BI model comes from eight operational systems at SUST, as shown in Fig. 2. Interviews with key personnel at each data source department were conducted to clarify understanding of the data. First, is the facilities management system that encompasses data about all facilities including class and lecture rooms, laboratories, libraries, offices, cafeterias, etc. Second is the enrollment system that contains students' demographic information and registration data. Third is the academic system from which all data related to courses and students' performance is acquired. Fourth system is the human resource system that covers educators' data. Fifth is the timetable system which comprises the schedules of lectures, seminars, etc. Sixth system is students' affairs which has all the financial support data. Seventh is the research system that contains data about the many research projects at SUST. Finally, the last system is the ICT support system that includes data about computers, internet connectivity, applications including the Learning Management System (LMS), and all technology support related data. The proposed BI framework integrates this carefully chosen data from all these sources into a central repository to be used by the appropriate stakeholders.

TABLE 27 DATA SOURCES

Data about	Factor	Operational systems
Campus, class and lecture rooms	Environment	Facilities management
Academic year, semester, student, college, department	Student Teaching Advising	Enrollment
Course, educator, course assignment, assessment method, references, mark, advisor, student GPA		Academic
Educator		HR
Timetable		Timetable
Financial support	HEI	Students affairs
Research		Research management
Computers, internet speed, WI-FI availability, LMS	ICT	IT management

Table 1 displays a mapping of the data sources details to their corresponding learning process factors.

In the second phase of the methodology, a DW is composed of subject-oriented *data marts* using one or more schemas to represent specific business or processes. It is obvious that Kimball's data marts are more appropriate for this model [16], as their staggered approach satisfies the business requirements of the framework and is yet flexible and scalable, i.e., it grows with the availability of additional data.

The DW was modeled and designed by applying the dimension modeling approach. Content analysis methods resulted in the identification of many factors influencing BI implementations. To meet the objective of the research, this case study implementation targets three learning process subject areas, namely, Course-Progress, Class-Occupancy, and Students-Registration. These processes cover the Environment, Teaching, Student and the HEI factors. Exploratory interviews with key stakeholders from the earlier identified SUST departments helped in highlighting and prioritizing these recognized factors. As well, KPIs, possible queries and information dashboard were created based on their feedback so as to study the possible correlation between the factors, their significance and thus validate the framework. Each of the three target subject areas is represented with a data mart that is composed of one central table with a multipart key called the *Fact* Table and a set of smaller related

tables called *Dimension* Tables. The data marts are:

1. *Course-Progress*

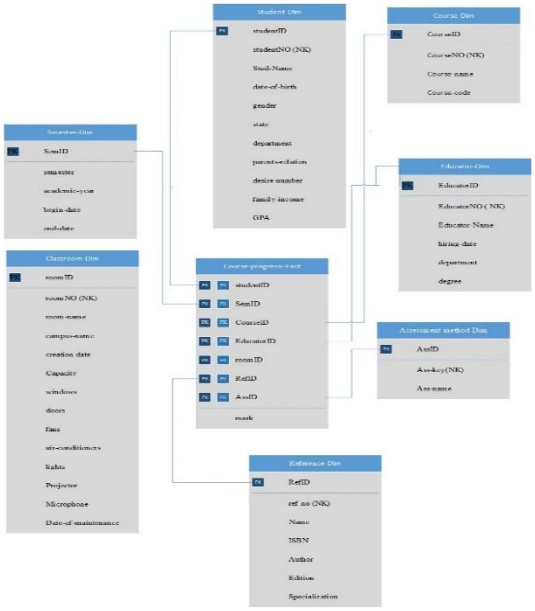


Fig. 1. Course-Progress data mart

One of the most important key indicators of a successful learning process is students' academic performance. So, building a data mart to analyze data related to students' and their grades is extremely helpful in tracking the *Teaching* factor. The fact table for this data mart is called Course-Progress as shown in Fig.3.

The dimension tables are: *Semester* that is composed of semester id, academic year, semester begin and end dates, etc. The *Course* dimension table contains data about courses such as Course id, name, code, etc., while the *Educator* dimension includes educator id, name, qualifications, hiring date, department, etc. *Student* dimension includes student's information such as student id, name, gender, date of birth, etc.*Reference* dimension table includes reference id, ISBN, Name, etc. *Assessment method*dimension table contains data about which assessment method is used. The last dimension table is *Classroom* and it has all the information about the classroom environment including class or lecture room id, name, campus, capacity, windows, doors, light, fans, air conditioner, projector or not, and microphone or not, etc.

2. *Class-Occupancy*

Managing class and lecture rooms is one of the big challenges facing HEIs, and has a huge impact on the learning process. It is quite helpful to examine which classrooms were used the most, and least, per academic year, and to explore which had the most, and least, equipment failures, etc. This information is vital for maintenance and load management as it relates to the *Environment* factor.

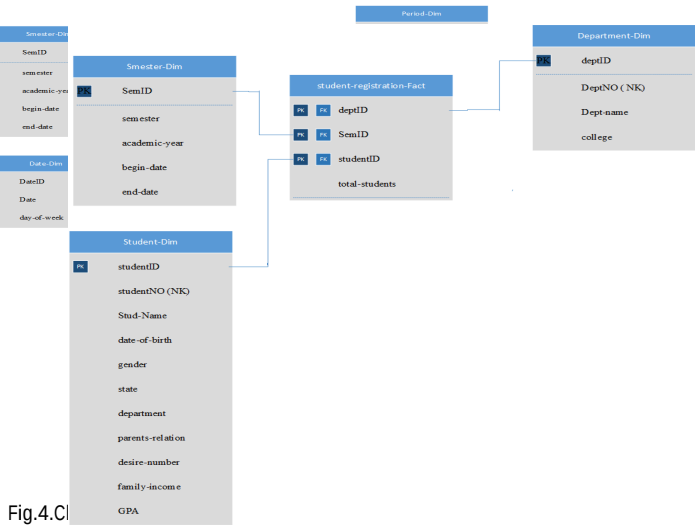


Fig.4.Cl

Fig.5. Students-Registration

The fact table in this data mart, shown in Fig. 4, just views the relationship between all its dimensions. This type of table is called *Factless* table, as it does not contain measures [16]. The dimension tables are: *Semester* and *Classroom*, and they are similar to the ones with the same names described in the previous Course-Progress data mart shown Fig.3. The *Date* dimension contains information about important dates in each semester such as date id, date, day of week such as Sunday, Monday, etc. The *Period* dimension table contains the time during the day, i.e., start and end times, for example 8 to 10, etc. The *Department* dimension includes, department id, name, college, etc. The *Occupancy* dimension contains data about occupancy status. The fact table for this data mart is *Classroom* and it consists of the keys of all the discussed dimension tables.

3. *Students-Registration*

Trends of students' enrollment numbers at colleges and departments over time are a strong indicator of the HEIs' competitiveness and success. This information is captured with the *HEI* and *Student* learning process factors discussed earlier. Thus, the BI model tracks students' registration and enrollment information per department over time to enable analyzing trends.

The fact table for this data mart is Student-Registration, shown in Fig. 5, and it contains keys of all dimension tables together with students' aggregates as measures. The dimension tables are *Semester*, *Department*, and *Student*. They are all

identical to the dimension tables with the same names described in the two previous data marts.

After laying out the methodological steps of the research approach, the next section presents the implementation of the three data marts, and discusses a selection of significant results. The purpose is to effectively expose the interrelationship between seminal factors and improvements in the learning process.

IV. DATA ANALYSIS

After completing the DW design, ETL procedures were carried, and the DW ready for the analytical phase. The open source BI tool, SpagoBI [17], was utilized for the model implementation to explore the benefits achieved by stakeholders from adopting BI technologies and their impact on the learning process at HEIs.

Information dashboards and reports were created using SpagoBI that reflect key stakeholders' perspectives. The knowledge extracted from the three data marts is discussed in detail, and its impact on the study objective is outlined.

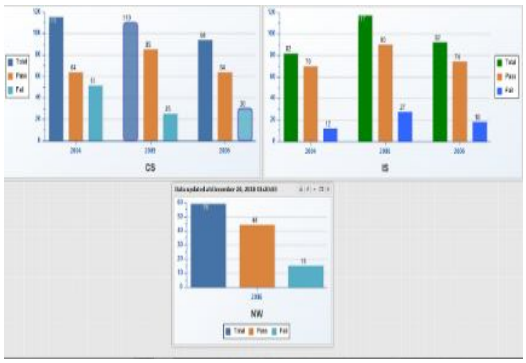


Fig.6.Students' performance

First, the Course-Progress data mart, avails students' performance data per course over time to key stakeholders such as college deans, department heads and educators. This academic performance information includes the total average, minimum, maximum, passing and fail students.

The dashboard shown in Fig. 6 contains information about the *Data Structures* course taught at the College of Computer Science and Information Technology. The course was taught at the three departments in the college, namely, Computer Science (CS), Information System (IS), and Networks (NW) during the 2004, 2005, and 2006 academic years in semester four. Enrollments in the NW department began in 2005, so there is data for the 2006 academic year only. The dashboard displays the total number of students registered for the course per department together with the number of passes and fails. The percentage of student who passed the course at the CS department was 77%, 55%, and 68% for the 2004, 2005, 2006 academic years respectively. Similarly, at the IS department the results were 85%, 78%, and 80% for the same period, and 74% at the NW department in 2006. It is clear that the number of students who passed the course at the IS department is greater than that of the two other departments. This information suggests that maybe there is a difference in the teaching environment, educator, assessment methods, students, etc., and that further investigation of this finding is warranted.

Additionally, from Fig. 6 we can see there was high failure among the CS department, 51 students, during the 2004 academic year. These finding lead the Dean and department heads to explore what made performance differ in departments while the course content is same. Correlating students' grades, educators, and assessment methods and any other information related to the *Teaching* factor might lead to some answers. The BI model fully supports and empowers these kinds of investigations.

Fig.7. Students' performance data

Fig. 7 shows a dashboard that was created from the Course-Progress data mart. It depicts the average, maximum and minimum grade for each department during the 2004, 2005, and 2006 academic years. The dashboard clearly shows that the average grades at the CS department for this period were 40%, 49%, 42% respectively. While they were 51%, 50%, 48% at the IS department for the same period, and 49% at the NW department in 2006. The fact that the average grade was consistently less than 50% strongly indicates that

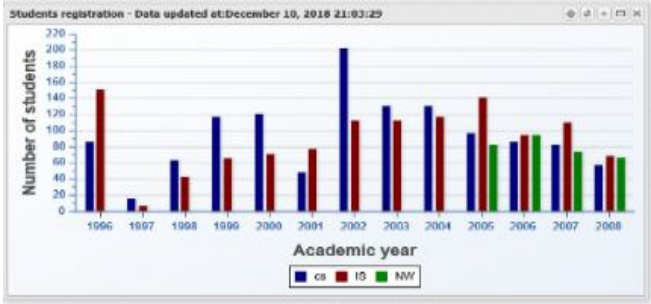


Fig.8. Firstyear students' registration

the overall students' performance was not good enough, and there is weakness in *one* or *more* factor that need improvement.

The third dashboard results, shown in Fig. 8, are from the Students-Registration data mart. It shows students' registration information for CS and IS departments between 1996 and 2008, and for the NW department from 2005 to 2008. These results show that in 1997 less than 20 students registered for both the CS and IS departments. Similarly, in 2002 the total number of students who registered for the CS department was about 200 and this is the largest number throughout the study period. The results also show that between years 2005 to 2008 the total number of registered students in the IS department wasconsistently greater than the CSdepartment.All of the information derived from this dashboard poses many questions that need answers. For example, why the number of CS and IS students in 1997 was so low? What is the reason behind the IS department accepting more students than the CS and NW departments during the 2005 to 2008 period? etc. The strength of the BI model is that it empowers such questions and inquiries while providing the tools and methods to analyze and find possible answers.

From the analysis above it is clear that using BI to investigate the many factors affecting the learning process, leads to information that can be used to determine weaknesses and to make improvement decisions.

V. CONCLUSION

This paper has investigated the adoption of BI technologies to enhance the learning process at HEIs. It has developed a basic BI implementation model, identified essential factors for its adoption, and empirically examined its efficiency. The case study implementation utilized as part of the validation, focused on four of the core factors identified; environment, teaching, student and HEI.

The analysis of the results obtained from the model has demonstrated the efficiency and benefit of using BI tools to analyze data and obtain valuable decision-making information. The implemented model is very robust, and additional data will lead to more information about factors, and the relationship between factors and learning process.

Furthermore, the study has laid out the foundation for improving the learning process while providing insight to HE planners, educators, and administrators to better realize the value of BI, the possible obstacles, and the existing leverage in its adoption. Future works involve extending the empirical study by involving additional data and factors as well as including HEIs at other regions.

II. REFERENCES

[1] Persson, Jacob, and Erik Sjöö. "Business Intelligence–its Impact on the Decision-Making Process at Higher Rducation Institutions: A Case Study at Karlstad University." (2017).

[2] Khan, Rafi Ahmad, and S. M. K. Quadri. "Business Intelligence: An Integrated Approach." *Business Intelligence Journal* 5, no. 1 (2012): 64-70.

[3] Betoret, Fernando Doménech, and Adela Descals Tomás. "Evaluation of the University Teaching/Learning Process for the Improvement of Quality in Higher Education." *Assessment & Evaluation in Higher Education* 28, no. 2 (2003): 165-178..

[4] Pavione, Caroline Stéffani Santos Nério, Bruna Camargos Avelino, and José Roberto de Souza Francisco. "Factors that Influence the Teaching-Learning Process from the Perspective of Accountancy Students: Analysis at a Higher Education Institution in Minas Gerais." *Revista de Educação e Pesquisa em Contabilidade* 10, no. 2 (2016).

[5] Wafa Ali Mohammed *"Investigating Factors Affecting the Learning Process in Higher Education - A Case Study"* International Journal of Computer Trends and Technology 68.6 (2020):10-17.

[6] Khan, Rafi Ahmad, and S. M. K. Quadri. "Business Intelligence: An Integrated Approach." *Business Intelligence Journal* 5, no. 1 (2012): 64-70. Power, Daniel J. "A Brief History of Decision Support Systems." *DSSResources. com* 3 (2007).

[7] Ranjan, Jayanthi. "Business Intelligence: Concepts, Components, Techniques and Benefits." *Journal of Theoretical and Applied Information Technology* 9, no. 1 (2009): 60-70.

[8] Chaudhuri, Surajit, Umeshwar Dayal, and Vivek Narasayya. "An Overview of Business Intelligence Technology." *Communications of the ACM* 54, no. 8 (2011): 88-98.

[9] Ta'a, Azman, Muhamad Shahbani Abu Bakar, and Abdul Razak Saleh. "Academic Business Intelligence System Development Using SAS® Tools." In Workshop on Data Collection System for PHLI-MOHE, vol. 13, p. 14.(2006).

[10] Kabakchieva, Dorina. "Business Intelligence Systems for Analyzing University Students Data." Cybernetics and Information Technologies 15, no. 1 (2015): 104-115.

[11] Piedade, Maria Beatriz, and Maribel Yasmina Santos. "Business Intelligence in Higher Education: Enhancing the Teaching-Learning Process with a SRM System." In 5th *Iberian Conference on Information Systems and Technologies*, pp. 1-5. IEEE, 2010. Dell2008

[12] Dell'Aquila, Carlo, Francesco Di Tria, Ezio Lefons, and Filippo Tangorra. "Business Intelligence Applications for University Decision Makers." *WSEAS Transactions on Computers* 7, no. 7 (2008):1010-1019.

[13] Aziz, Azwa Abdul, Julaily Aida Jusoh, Hasni Hassan, Wan Mohd Rizhan Wan Idris, Addy Putra MD Zulkifli, and Sharul Anuwar Mohamed Yusof. "A Framework for Educational Data Warehouse (EDW) Architecture Using Business Intelligence (BI) Technologies." *Journal of Theoretical & Applied Information Technology* 69, no. 1 (2014).

[14] Aziz, Azwa Abdul, W. M. Rizhan, W. Idris, Hasni Hassan, and Julaily Aida Jusoh. "Intelligent System for Personalizing Students' Academic Behaviors-A Conceptual Framework." International Journal on New Computer Architectures and Their Applications 2, no. 1 (2012): 138-153.

[15] Izzeldin A. Elhassan *"Investigating the Success of Business Intelligence in Aligning Higher Education and Labor Market"* International Journal of Computer Trends and Technology 67.12 (2019):62-70.

[16] Kimball, Ralph, and Margy Ross. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. John Wiley & Sons, 2011.
[17] Gioia, Andrea, Grazia Cazzin, and Ernesto Damiani. "SpagoBI: A Distinctive Approach in Open Source Business Intelligence." In 2008 2nd IEEE International Conference on Digital Ecosystems and Technologies, pp.592-595. IEEE, 2008.