

K-12 Students' Academic Status: A Data Warehouse Architecture Framework

Evangeline T. Sarte, Thelma D. Palaoag

Abstract— *The K-12 Basic Education Program in the Philippine is now on its full swing status. However, as it moves to the peak of its full implementation, the number of drop-outs, retention, and migration students is increasing. With this in mind, the researcher came across with designing data warehouse and data mining architecture in the analysis of drop-out, retention, and migration patterns of students. Academic performance is the main factor in the students' drop-out, retention, and migration. The proposed architecture would be sufficient for the analysis of the K-12 students' academic status. It served as a foundation in the conduct of a thorough study on students' drop-out, retention, and migration patterns.*

Keywords— K-12, data warehouse architecture, academic status.

I. INTRODUCTION

In the past decades, there have been significant reforms in education systems worldwide, one of which is the adoption of the K-12 Basic Education Program in the Philippines.

For years, the Philippines have been lagging behind other developing countries in the ASEAN region. It still adopts the K-10 system, until 2011, when the government has implemented the K-12 educational system in the Philippines to get along with the educational system of other countries all over the world.

However, K-12 brought confusion in the society not only to the parents and students but also to the academe. Though K-12 is in full swing in the Philippines for eight years, confusion still exists in the mind of the parents and students and sometimes the reason for drop-outs, migration, and retention of students. According to [1], there is a tremendous increase in the number of elementary and high school drop-outs. Since 2012, it reached 4.8 million or an 11 percent increase.

One of the dilemmas of decision-makers in the educational sector is the increasing number of drop out of students. There is always a practical reason why students were dropping out of school. The increasing rate of student drop-outs every year is a serious issue that the Department of Education and other education sectors all over the world should solve. It might result in a high rate of unemployment, which can give a significant impact on the country's progress [2].

According to [3], dropping out is the result of students' years of academic difficulties, mistakes, and wrong choices. Some students' drop-out because of family responsibility, they need to help their parents to take good care of their

siblings and others to help sustain the financial problem faced by the family. Moreover, it is also because the students got bored and cannot relay the academic connection in real life. It is also because students cannot mingle to his classmates and teachers, an inferiority type of feeling. In general, it is because of emotional and academic reasons why students' dropped. The reasons were given by students vary, but the results and consequences of their decision to drop-out is similar, and that is repeating the grade level on the next time they enroll.

Another focus of school administrators is the increasing number of repeater senior high school students. They are students who need to enroll in the same year level again [2]. In this research, retention was used by the researcher to define the repeaters or the year retainers. In 2013, repeaters in secondary education in the Philippines were 2.3% [4]. The increasing number of student repeaters is really disturbing, and it requires further attention and analysis.

In today's education sector status, secondary education schools also aimed to improve the quality of the Senior High School education. At the same time, the students' academic performance success because the ability to meet students' needs represents the quality of their teaching process. The quality of services they provide should meet the needs of the students in the educational process.

Many researchers have come up with research investigating the factors related to the students' dropout, retention, and migration. All secondary schools and higher education institutions in the Philippines were looking into how this could be prevented to happen on its early stage.

Meanwhile, a data warehouse is a combination of data from different sources. It served as the business source of truth used in decision making. It provides analytical reports which help decision-makers in the decision process [7]. It acts as a decision support system that provides the information required by the decision-makers [8].

In the education domain, there are studies on data warehouse development, one of which is the study of [9] wherein he summarized the process on how to design and model an academic data warehouse as well as how the existing institution facilities and databases could support decisional activities. Likewise, [10] have specified the opportunities associated with the development of a data warehouse for academic institutions.

Revised Version Manuscript Received on 10 September, 2019.

Evangeline T. Sarte, Our Lady of Pillar College-Cauayan, Cauayan City, Philippines. (Email: vansarte@yahoo.com)

Thelma D. Palaoag, University of the Cordilleras, Baguio City, Philippines. (Email: tpalaoag@gmail.com).

On the other hand, data mining educational information is critical as it explores unique types of data from educational data, files, and reports [11]. Also, it aims to obtain a better understanding of students' learning processes and to come up with better methods and process to improve the quality of educational system offered by institutions [12].

There are data mining algorithms available that decision-makers can use to achieve their desired results. These algorithms can perform data exploration and extraction. The well-known algorithm is the clustering algorithms wherein it can group data into different consistent n- groups and recognizes patterns. The results can help decision-makers in the creation of a better decision model [13].

Likewise, Educational Data Mining (EDM) focuses on essential mining patterns and knowledge discovery from the educational sectors' information systems and applications. It aimed to discern valuable knowledge to help the educational institutes manage and improve their students' academic performances, deliverables, and status [14].

The study aimed to develop a data warehouse and data mining architectural framework useful in the analysis of students' drop-out, retention, and migration patterns. This study would help the administrators to improve their strategic plans that could reduce or eliminate students' drop-out and migration and also, to find ways on how to avoid retention of their students on the same grade level.

II. METHODOLOGY

The researcher applied multiple methods of data gathering such as stakeholders interview on their expectations; document, forms, and reports analysis; and analysis of the institutions' operational process. In the personal interview, the researcher was able to identify the students' demographic profile. Operational process input and output investigation were conducted to determine that data being process by the different schools and the expected output required. In the conduct of the document analysis, the researcher reviews existing data warehouse and data mining architecture for educational institutions proposed by other researchers which serves as the foundation in the development process of a K-12 students' academic status data warehouse and data mining architecture. While in the stakeholders' interview, the expectations for designing data warehouse and data mining solution was determined as well as the expected features and functionality of the data warehouse and data mining.

III. RESULTS & DISCUSSION

Through an interview with the different school principals, the researcher found out that reports on student's drop-out, retention, and migration were made and consolidated by the school Registrar. Based on their existing business process, retrieving reports happen in the school registrar and school principal's desktop. Report information includes student number, student name, address, contact number, parent's name, family income, status (whether drop-out, grade retention, change track, and transferred out), and remarks.

The data warehouse is a central repository capable of extracting data from different data sources. Through ETL (Extract, Transform, and Load) process that the data

warehouse can perform, the end-users were guaranteed a correct format, cleansed, and quality guaranteed data output. The data output was released only when it is suitable for stakeholder's consumption [15] [16]. It deals with a detailed level of data and digests data. Wherein data can be current and historical data used in the creation of reports needed in the decision-making process [17]. The researcher based the proposed data warehouse architecture from the data warehouse architecture of [17] [18]. The researcher perceived that the architecture applied to the K-12 'students'

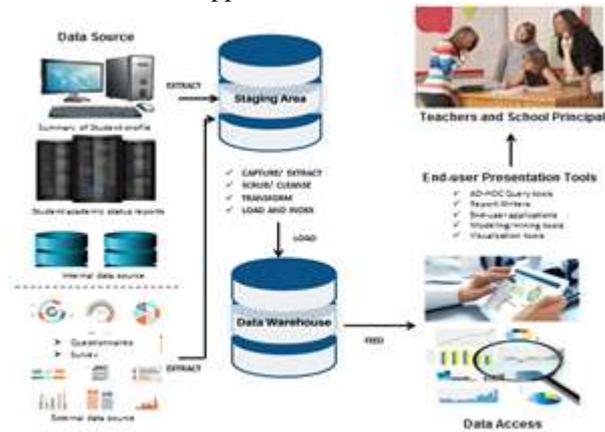


Figure 1. K-12 Students' Academic Status data warehouse architecture

academic status could help in the transformation of an operational database into an informational warehouse. It is beneficial for decision-makers to conduct data analysis, prediction, and forecasting using the proposed data warehouse architectural framework shown in Figure 1.

The architecture presents the different phases in transforming educational institutions' operational databases into a data warehouse. It includes the internal/external data source section, the staging area section, the data warehouse section, the data access section, and the end-users. The process of data capture and extraction, data scrubbing and cleansing, data transformation, and data loading and indexing happen in the staging area section.

Based on the architecture, the summary of students' profiles, students' academic status reports, and output of surveys and questionnaires serve as the data source. The student demographics and information were extracted and stored temporarily into the buffer. Then it is pre-processed and moves forward to the staging area. In the staging area, it is where the data cleansing; data scrubbing and reconciliation; fixing data entry errors; and normalization of data happens. After the data normalization process, the converted data were loaded and indexed into the data warehouse. Figure 2 depicts the staging area process in the data warehouse.

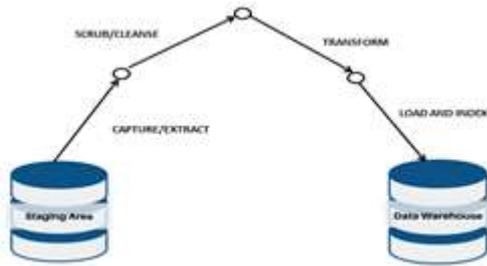


Figure 2. Staging area process in the data warehouse

Only student status-related items retained in the data warehouse in the process of dropping tables. Merging of tables also happens, which results in the creation of new tables. Also, new columns could be added or removed in the transformation process. Cleaned and transformed data are transferred to the newly built data warehouse.

Data Mining is a process that is useful for searching hidden facts from data coming from the different students' information systems. It describes the data from a different source, and it converts into meaningful information. Pre-processing is the most required method to describe the data sets, which removes unnecessary information from the collected data and retains the relevant information. It converts all the attributes to its category [19]. Figure 3 shows the Data mining architecture [20].

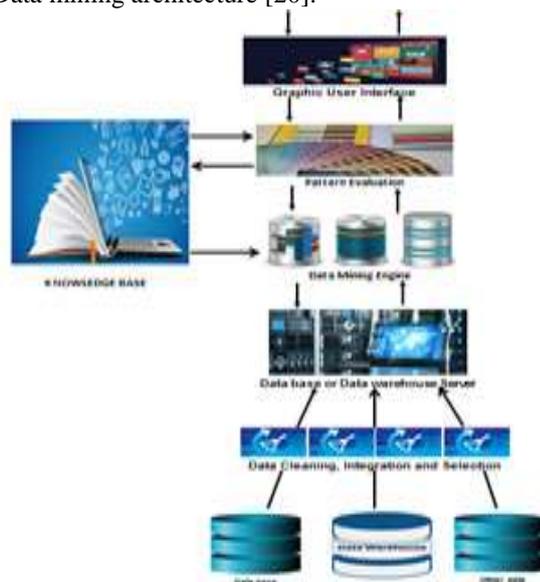


Figure3. Data mining architecture

Data mining should be appropriately implemented to determine student academic status patterns and trends, which is essential to school administrators as they perform their decision making.

The data source is the actual source of data that comes from the school database, data warehouse, or other data repositories. It undergoes the process of data cleaning, integration, and selection; and then stored in a data warehouse server.

Aside from the server contains the actual data to be processed, it performs data retrieval based on the data mining algorithm used by the end-users. Along with it, is the data mining engine which is the essential element of the data mining architectures as it is used to perform data mining tasks such as association, classification, characterization,

clustering, and prediction.

Furthermore, the pattern evaluation module is mainly focusing on the measurement of the integrity of the pattern. A starting value could be used to measure its integrity. The data mining engine searches for exciting patterns, and it interacts with the pattern evaluation module.

Besides, the graphic user interface (GUI) module is used to interconnect between the stakeholders and the data mining system. This module helps the stakeholders use the system quickly and efficiently. In times when they specify a query, this module works together with the data mining system, and then the result is displayed comprehensively.

Lastly, the knowledge base module serves as a guide in searching for the result patterns. It contains user beliefs and user data, which is necessary for the data mining process. The data mining engine gets inputs from this module to come up with a more accurate and reliable output. Regularly, the pattern evaluation module interrelates with the module for data input and update.

IV. CONCLUSIONS

The data warehouse architecture would help school decision-makers in performing student academic status and record analysis. It also helps them determine the critical patterns and trends of students who might drop-out, change track, grade repeater, or transferred out.

The educational data mining is also an essential method that can be used to extract remarkable, interpretable, valuable, and unique information about student academic status. It could lead to a better appreciation of tools that can be used to determine students' drop-out, retention, and migration patterns.

REFERENCES

1. D. Porcalla (2017), Number of elementary, high school dropouts rising- lawmakers, The Philippine Star, available at <https://www.philstar.com/headlines/2017/06/25/1713711/number-elementary-high-school-dropout-s-rising-lawmaker>.
2. K. S. Benedicto, M.E. Ochea, R.J. Reyes, Problems Encountered By Dropouts, Repeaters, Balik-Arals Andstudents With Failed Grades In Public High Schools Inquezon City: A Basis For A Proposed Guidance intervention Program , Our Lady of Fatima University, Quezon City, unpublished.
3. R. Furger (2011), How to End the Dropout Crisis: Ten Strategies for Student Retention, Dropout Prevention, George Lucas Educational Foundation.
4. World Data Atlas (2013), Philippines-Repeaters in Secondary Education, available at <https://knoema.com/atlas/Philippines/topics/Education/Secondary-Education/Repeaters-in-secondary-education>
5. J. Mateo (2018), Student population soars to 'historic' 27.7 M, available at <https://www.philstar.com/other-sections/education-and-ome/2018/05/31/1820117/student-population-soars-historic-277-m#kwFDH4bqOjjiu7eo.99>
6. P. Daluz (2018), 3 Things to consider before changing strands in senior high, available at <https://blog.edukasyon.ph/senior-high/3-things-to-consid>

- er-before-changing-strands-in-senior-high/
7. L. Wainstein (2018), Data Warehouse Architecture – An Overview, available at <https://techburst.io/data-warehouse-architecture-an-overview-2b89287b6071>
 8. TN. Manjunath, R. Hegadi & Umesh, & G K, Ravikumar. (2011). Design and Analysis of DWH and BI in Education Domain. International Journal of Computer Science Issues. 8.
 9. C. Dell'aquila, 'An Academic Data Warehouse' World Scientific and Engineering Academy and Society (WSEAS) Stevens Point, Wisconsin, USA ©2007.
 10. McMillen and R. McBroom, 'what academia can gain from building a data warehouse' No.1, pp.41-46.
 11. C. Romero and S. Ventura, Educational Data Mining: A Survey from 1995 to 2005, Expert Systems with Applications, 33(1), 135-146, 2007.
 12. M. Zorrilla, D. Garcia, and E. Alvarez, A Decision Support System to Improve eLearning Environments, In Proceedings of the 2010 EDBT/ICDT Workshops, DOI:10.1145/1754239.1754252, 2010.
 13. A. Algarni, (2016). Data Mining in Education. International Journal of Advanced Computer Science and Applications. 7. 10.14569/IJACSA.2016.070659.
 14. A. A. Saa (2016), Educational Data Mining and Students Performance Prediction, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 5, 2016.
 15. M. Malipot (2018), Its back to School for 28M Students, Manila Bulletin. Available at <https://news.mb.com.ph/2018/05/27/its-back-to-school-for-28-m-students/>
 16. B. Gumba (2016), Gender Equality in Education and Student Participation in Higher Education Institution: A Case in the Philippines, IJSRST, Volume 2, Issue 6, Print ISSN: 2395-6011, Online ISSN: 2395-602X, Available at https://www.academia.edu/31221469/Gender_Equality_In_Education_And_Student_Participation_In_A_Higher_Education_Institution_A_Case_In_The_Philippines
 17. P. A. San Buenaventura (2019), Education Equality in the Philippines, International Workshop on Data Disaggregation for the Sustainable Development Goals, Available at https://unstats.un.org/sdgs/files/meetings/sdg-inter-workshop-jan-2019/Session%2011.b.3_Philippines_Education%20Equality%20Assessment_FINAL4.pdf
 18. C. Ensoy and M. Lavega (2015), High School General Point Average (GPA) as Basis for Students' Academic Performance in College, JPAIR Institutional Research, Vol. 5, No. 1, DOI: <https://doi.org/10.7719/irj.v5i1.344>
 19. Philippine Statistics Authority (2013), The Educational Attainment of the Household Population, Population and Housing, Available at: <https://psa.gov.ph/content/educational-attainment-household-population-results-2010-census>
 20. Philippine Statistics Authority (2018), Births in the Philippines (2017), Available at: <https://psa.gov.ph/content/births-philippines-2017>
 21. B. Balamban and M. Addawe (2016), Poverty incidence among Filipinos registered at 26.3% , as of first semester of 2015, Philippine Statistics Authority, Available at https://psa.gov.ph/sites/default/files/attachments/ird/pressrelease/Press%20Release_poverty.pdf
 22. World Data Atlas (2013), Philippines- Repeaters in Secondary Education, Available at <https://knoema.com/atlas/Philippines/topics/Education/Secondary-Education/Repeaters-in-secondary-education>
 23. C. Bennet (2017), Daily School Attendance Matters! Negative Impact of Absenteeism for all Grades and Socio-Economic Groups, Available at: <https://www.thoughtco.com/daily-school-attendance-matters-4084888>
 24. A. Latif, A.I. Choudhary, A. A. Hammayun, (2015), Economic Effects of Student Dropouts: A Comparative Study, J Glob Econ 2014, Vol 3(2): 137, DOI: 10.4172/2375-4389.1000137
 25. L. Narra, T. Sahama, and P. Stapleton, "Clinical Data Warehousing A Business Analytics approach for managing health data", Proceedings of the 8th Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2015), Pages 101 104, Sydney, Australia, January 2015.
 26. R. Kimball and M. Ross, "The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling", Third Edition, John Wiley & Sons, Inc., Indiana, United States of America, 2013.
 27. T. M. J. Al Taleb and M.M. K. AL Saedi (2017), Data Warehouse System for Higher Education Student Information System, IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 19, Issue 5, Ver. III (Sep.- Oct. 2017), PP 47-53
 28. Y. Bassil (2012), A data warehouse design for a typical university information system, Journal of Computer Science & Research (JCSCR) - ISSN 2227-328X, Vol. 1, No. 6, Pages. 12-17, December 2012
 29. N. Priyadarshini and M. Ray (2017), A review: Data mining techniques in Education Academia, ISEC 2017 workshops: ModSym, DIAS and EDUDM, Jaipur, India.
 30. Dataflair Team (2018), Data Mining Architecture-Data Mining types and techniques, available at: <https://data-flair.training/blogs/data-mining-architecture/>.