Profiling Oman education data using data visualization technique

Sultan Juma Sultan Alalawi, Izwan Nizal Mohd Shaharanee, and Jastini Mohd Jamil

Citation: AIP Conference Proceedings **1782**, 020003 (2016); doi: 10.1063/1.4966057
View online: http://dx.doi.org/10.1063/1.4966057
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1782?ver=pdfcov
Published by the AIP Publishing

Profiling Oman Education Data using Data Visualization Technique

Sultan Juma Sultan Alalawi[1, a)], Izwan Nizal Mohd Shaharanee[1, b)] and Jastini Mohd Jamil [1, c)]

[1]*School of Quantitative Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*

[a)]Corresponding author: sultanjuma44@gmail.com
[b)] nizal@uum.edu.my
[c)]jastini@uum.edu.my

Abstract. This research works presents an innovative data visualization technique to understand and visualize the information of Oman's education data generated from the Ministry of Education Oman "Educational Portal". The Ministry of Education in Sultanate of Oman have huge databases contains massive information. The volume of data in the database increase yearly as many students, teachers and employees enter into the database. The task for discovering and analyzing these vast volumes of data becomes increasingly difficult. Information visualization and data mining offer a better ways in dealing with large volume of information. In this paper, an innovative information visualization technique is developed to visualize the complex multidimensional educational data. Microsoft Excel Dashboard, Visual Basic Application (VBA) and Pivot Table are utilized to visualize the data. Findings from the summarization of the data are presented, and it is argued that information visualization can help related stakeholders to become aware of hidden and interesting information from large amount of data drowning in their educational portal.

INTRODUCTION

Before the discovery of oil, the development of education system in Sultanate of Oman is progressing slowly. The discovery of oil had contributed to a better education system in Sultanate of Oman. In 2007, the Sultanate of Oman inaugurated Information System for Ministry of Education called "Educational Portal". The system developed to support administration of studies and on-line learning concepts. Thousands of employees, teachers, students and student parents use it for many activities daily. It stores educational data, which includes of all information about students and their interaction with teachers and the employee of ministry. It also delivers online materials to students. It supports faculties in examinations, semi-automatic evaluation of electronic exams or skimmed written exams, training, theses, publications, discussion groups, assessment and making the results of capacity-to-study tests public. The Educational Portal provides multi-functionality for various external systems. Furthermore, the database of Educational Portal contains history of every user's interactions.

In spite of its promise, managing data does have challenges and risks. Data drown into education databases and data warehouses faster than it can be managed and very difficult to analyze. Data can be meaningful only if it can provide insights into its. Institutions should deploy the suitable technology systems, techniques and tools that can assist and make this possible. With the emergence of technology and massive data generation, data mining offer a ways for proper and systematics analysis and clarity into the data.

Data mining defined as technique tools of extraction and exploitation unseen predictive information from huge data [1]. These powerful new technologies provide potential tools to assist the ministry of education in exploiting their data warehouses to produce useful information. Data mining applications allow the Ministry of Education to predict future trends and behaviors of student performance. This approach are able to discover underlying patterns in order to

The 4th International Conference on Quantitative Sciences and Its Applications (ICOQSIA 2016)
AIP Conf. Proc. 1782, 020003-1–020003-5; doi: 10.1063/1.4966057
Published by AIP Publishing. 978-0-7354-1444-0/$30.00

predict pupil outcomes for example needing extra help, dropping out, or being capable of more demanding tasks. The educational data mining and learning analytics have the good potential to make visible data for any education institution that have heretofore gone hidden, unknown, and as a result actionable. Data mining offer a quicker ways in answering complex queries that consuming more time to resolve [2]. For instance, how can students pick to usage educational software, it is sensible to concurrently consider data at the keystroke level, classroom level, session level, student level, and school level. One of the major tasks in data mining is interpretation and evaluation. These tasks offer the researchers to communicate the results and patterns clearly and efficiently via statistical graphics, plot and information graphics. Data visualization as defined by many researchers plays an important roles in representing data and information into certain format that easy to understand and it enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns.

This paper is organized into the following: Section II define the problem statement of this paper. In Section III illustrates the research work that has been conducted in data visualization and data mining. In Section IV consists of the description of the proposed framework for developing interactive data visualization using Microsoft Excel dashboard for Oman Education data. The conclusions and further work is discussed in Section V.

PROBLEM STATEMENTS

Ministry of education in Sultanate of Oman have huge databases contains massive information. These databases expand every year as many students, teachers and employees enter into the database. For every transaction, processes and activities occurred within the education system, a tremendous amount of data generated daily. It is stores massive data generated dynamically by student interactive with system like homework and discussion. Teachers also generated massive data in the system by their work in student report, result and exam. These data can offer new insights and provides great opportunities to make better evaluation of students' performance and behaviors. It can also enable faculty staff to taken to improve and enhance teaching methods, techniques and quality of offered study materials [3]. Exploring knowledge from this data increasing every day and from different sources is the biggest challenge. These data can be a valuable resource if used suitable analyze technique.

Currently, the Ministry of Education in Sultanate of Oman relies on the classical descriptive statistic to make analyze of their data. The statistic department utilized spreadsheet to account all information like (number of schools, owner of school, gender of schools, number of students, gender of students, student age, number of employees, etc.). With this classical approach, they faced big challenges to extract hidden information from huge amount of data. The decision maker often wants to access more data because they believe that by having large amount of data they can perfectly explain and interpret any phenomenon they faced. However, many data generated by the user are lacking in quality. Problem such as missing values, error in data and the existence of outlier offer occurred in large and huge databases. Thus, a suitable technique is needed in order to extract useful information from this data. Data visualization offers a way in the context of summarizing massive data into useful and easy to understand format.

RELATED WORKS

Many researches have been done in exploring student academic data using data mining techniques. [3] conducted a study on student success by collecting data from Electrical Engineering and Computing in academic. The data contain student record and final grades at first semester course. The main objective of this study is to investigate how data mining techniques can be used in the academic community to potentially improve some aspects of education quality. In this study researchers used WEKA software to analyzed student data. The outcome of their results indicated that data mining techniques definitely have a greater role in education environment, and should be used to improve education quality.

[4] presented a case study on educational data mining to investigation the most relevant subset features with minimum cardinality for achieving high predictive performance and evaluation of goodness of subsets with different cardinalities. The data of this study collected from Tamil Nadu, India, contain demographic details, family details, socio-economic details, previous academic performance at secondary level from different schools and other environmental details. The data obtained from students through a questionnaire with close-end questions, a total of student participants in study was 1969. Researchers applied Naïve-Bayes algorithm. The study revealed that increase in the predictive accuracy with the existence of minimum number of features.

[5] attempt to study the main attributes that may affect the student performance and how can enhancing the quality of the higher educational system. The data of this study was collected using a questionnaire passed among undergraduate students who took the Programming I course (C++). Researchers used WEKA software to analyze the data. Researchers build the classification model using the decision tree algorithm. The study showed that decision makers could enhance the student performance and educational system by utilizing the extracted knowledge.

The research by [2] has used decision tree algorithm as data mining methods to analyze student data. The main objective of this research is to develop a faith on Data Mining techniques by using education data. Student dataset was taken from community college database. It is contain records of 2000 student performance details in 4 attributes. The attributes are MAT score, verbal ability score, quantitative ability score and likelihood of placement. The researchers discovered that student performance and overall personality improvement could be more efficiency by exploiting Data Mining technique.

PROPOSED FRAMEWORK

The main objective of this paper is to visualize the data that will be utilized later for data mining tasks. These will offer an initial summary the characteristics of Ministry of Education in Sultanate of Oman data.

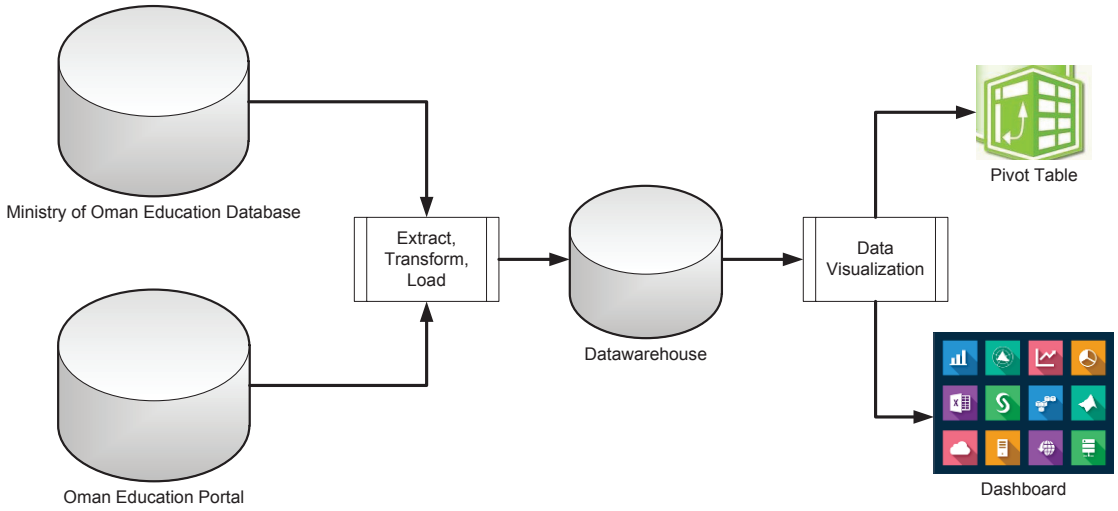

FIGURE 1. The proposed framework for Oman Education Information Visualization

1. Data Selection
   Students' data were collected from the Educational Portal system and the total number of students was 42499 students. The collected data was extracted and organized in Microsoft Excel format. Each student record had the following attributes: student ID, student name, school code, region code, school type, student gender, date of birth, nationality, and total of result in final exam.

2. Data Visualization Tools
   Excel Dashboards tools are utilized for data visualization. Excel Dashboard is a dashboard interaction paradigm has been chosen as it provides large amounts of information in one interface [6]. Dashboards compile key metrics in a simple and easy to interpret interface so that school officials can quickly and visually see how the organization is doing. Administrators automatically update dashboards based on data stored in student information systems. Software combines data from various streams to present a clear and comprehensive overview of school operations [7]. Dashboards provide quick access to key performance information that decision makers need to make effective decisions. The most important purpose of dashboards

tool is to represent the information in graphical format using the process of data mining [8]. Dashboard is a great way to summarize, analyze, explore, and present data to provide good understanding for decision makers.

3. Dashboard Development

There are three levels in developing Excel Dashboard, which are mentioned below:

Level 1: Data needs to been imported and connected to Excel Spreadsheet. Connecting this data into Microsoft Excel will create an easy and up to date data repository for information summarization. Figure 2 is a snapshot of the student data utilized in this research work.

Birthdate	Email	TownCode	ReligionID	GenderID	NationalityID	SchoolCode	Class	GradeID	StageID	EduSysID	SchoolT	School	ZoneID
29/09/1993	NULL	10087	1	1	512	5050	1	2247	936	4	3	1	30
21/09/1997	NULL	10004	1	2	1044	5054	1	2247	936	4	3	1	30
01/01/1987	NULL	10018	1	2	1008	5054	1	2246	936	4	3	1	30
05/09/1964	NULL	10091	1	2	1135	5054	1	2246	936	4	3	1	30
01/01/1984	NULL	10004	1	2	512	5054	1	2247	936	4	3	1	30
01/01/1966	NULL	103372	1	2	1044	5054	1	2247	936	4	3	1	30
01/01/1975	NULL	10003	1	2	1044	5054	1	2247	936	4	3	1	30
25/08/2003	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
15/10/1997	NULL	10004	1	2	512	5054	1	2247	936	4	3	1	30
10/06/2001	NULL	10004	1	2	512	5054	1	2247	936	4	3	1	30
20/03/1999	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
23/02/1961	NULL	10004	1	2	512	5054	1	2247	936	4	3	1	30
01/01/1965	NULL	10003	1	2	1044	5054	1	2246	936	4	3	1	30
25/06/1998	NULL	10003	1	2	1044	5054	1	2246	936	4	3	1	30
01/01/1971	NULL	10003	1	2	1044	5054	1	2246	936	4	3	1	30
01/01/1971	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
09/05/1959	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
01/01/1964	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
22/04/1962	NULL	10004	1	2	512	5054	1	2246	936	4	3	1	30
25/11/1964	NULL	10006	1	2	1044	5054	1	2246	936	4	3	1	30

FIGURE 2. A snapshot of Oman Education data extracted from Educational Portal

Level 2: In level two, pivot table is utilized to deal with large amount of data. Pivot tables are one of Excel's most powerful features. A pivot table allows you to extract the significance from a large, detailed data set.

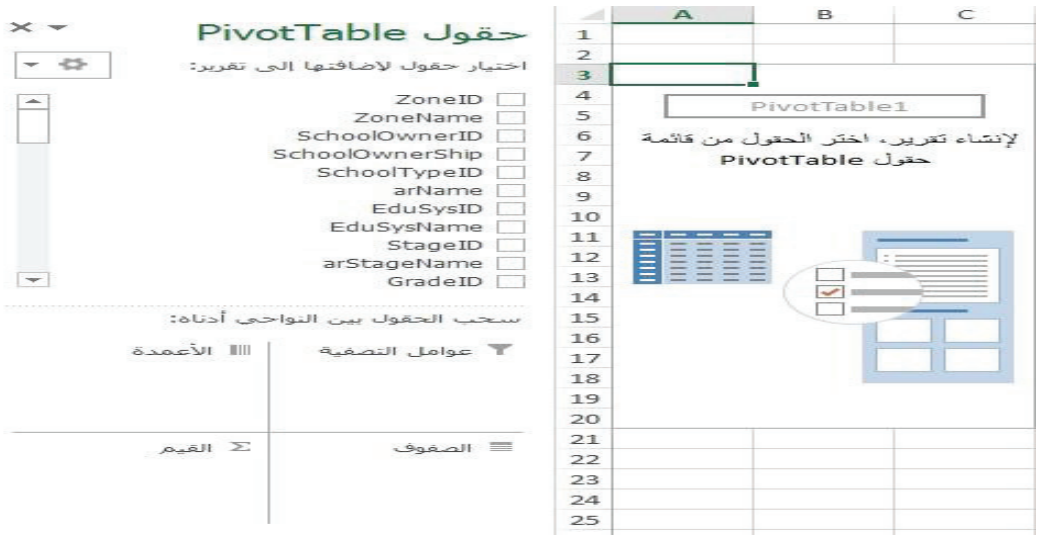

FIGURE 3. A snapshot of pivot table for Oman Education Dashboard

Level 3: To create the dashboard chart, the user can use a Pivot Chart to show dynamic visual reports. Pivot Chart can be created directly from your completed Pivot Table, making the chart creation process a snap.

Dashboard focus on how the data is being display and in what forms can have data visible to us. It lets users and decision makers to group and filter data in different ways so you can more easily make comparisons.

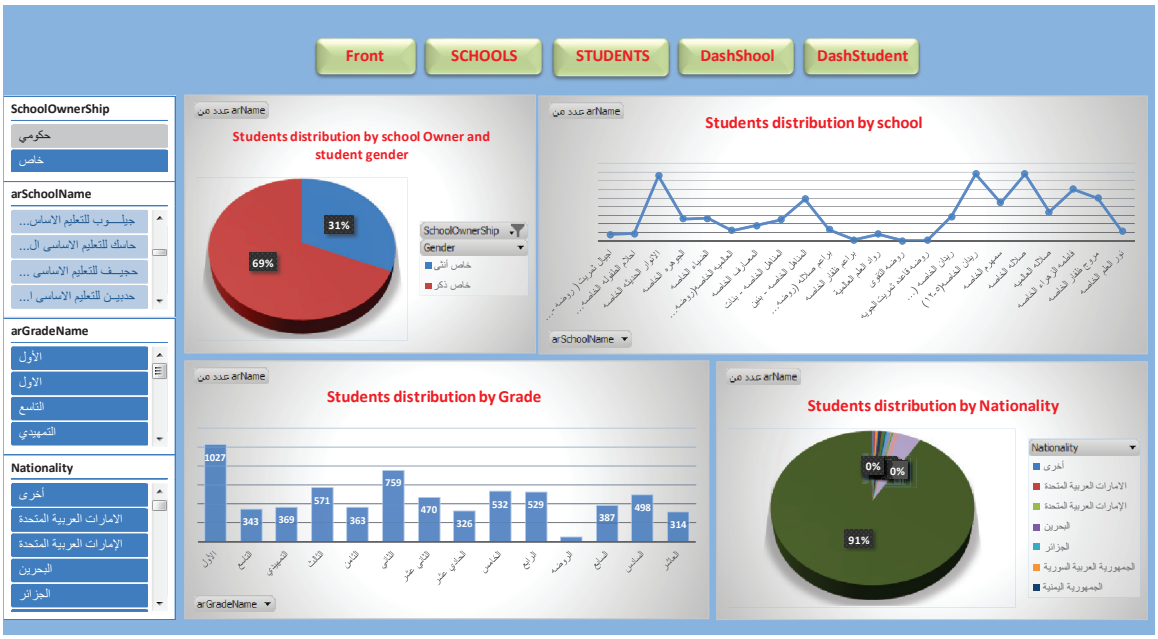

FIGURE 4. A snapshot of Oman Education Dashboard System

CONCLUSION

This research utilizes Microsoft Excel dashboard to visualize and summarizing student data into simple and easy to comprehend information. The data mining visualization can acts as a tools for mining hidden and useful information and knowledge from the extracted databases. This knowledge can be used to give a deeper understanding of student's descriptions and behaviors in the class under study. In addition, the faculty and managerial decision maker can utilize this innovative tool in making better and faster decision as this tools offer an integrated and easy-to-use graphical user interface.

REFERENCES

1. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco: Morgan Kaufmann Publishers, 2011.
2. S. Agarwal, G. N. Pandey, and M. D. Tiwari, "Data Mining in Education : Data Classification and Decision Tree Approach," *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 2, no. 2, pp. 140–145 (2012).
3. M. Vranic, D. Pintar, and Z. Skocir, "The use of data mining in education environment," *The 9th Int. Conf. Telecommun*ication, pp. 243–250, 2007.
4. M. Ramaswami and R. Bhaskaran, "A Study on Feature Selection Techniques in Educational Data Mining," *Journal of Computing*, vol. 1, no. 1, pp. 7–11, (2009).
5. Q. A. Al-Radaideh, E. M. Al-Shawakfa, and M. I. Al-Najjar, "Mining Student Data Using Decision Trees," *International Arab Conference on Information Technology* (ACIT, 2009)
6. S. Mazumdar, A. Varga, V. Lanfranchi, D. Petrelli, and F. Ciravegna, "A Knowledge Dashboard for Manufacturing Industries," in *The Semantic Web: ESWC 2011 Workshops: ESWC 2011 Workshops, Heraklion*, Greece, May 29-30, 2011, Revised Selected Papers, R. Garc'ia-Castro, D. Fensel, and G. Antoniou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 112–124.
7. D. West, "Big Data for Education: Data Mining, Data Analytics, and Web Dashboards," *Government Studies Brookings*, (2012).
8. V. Abrol and J. Malhotra, "Data Dashboard-Integrating Data Mining with Data Deduplication," *International Journal of Computer Applications*. 28-33, 71(2013).