Indonesian Journal of Electrical Engineering and Computer Science
Vol. 2, No. 2, May 2016, pp. 469 ~ 477
DOI: 10.11591/ijeecs.v2.i2.pp469-477

Development of a Decision-Making System for Sultan Moulay Slimane University in Beni Mellal, Morocco

Abdellah Amine*, Rachid Ait Daoud, Belaid Bouikhalene
Sultan Moulay Slimane University, Faculty of Science and Technology
Po Box 523, Beni Mellal, Morocco
Phone: + 212 (0) 661748520 Fax: +212 (0) 523 481351
*Corresponding author, e-mail: a.amine@usms.ma

Abstract

The issue dealt with in this article is to develop a decision-making information system related to the digital environment of the University work. We propose to model the data within the university in order to transform a system of information into a decision-making information system,that is based on thetrades databasesoriented toward the actors. A decision-making information is a system that allows the decision makers of the university to have relevant information and powerful analytical tools to help them take the right decision at the right time.

Keywords: *Meta data, strategic information systems, user classification model, data warehouse, Meta modelling, Sql Server*

1. Introduction

Decision-making and its execution constitute the fundamental purposes of any organization and any management. Indeed, any organization depends structurally on the nature of the decisions which are taken by its decision-makers. The decision-makers are more and more confronted with new situations and with an environment which quickly evolves. In such a context, decision-making has become extremely difficult. So, the intuitions and experience of the decision-makers are no longer sufficient for making the right decisions. In fact, the use of decision support systems is vital to determine the congruent information for decision-making. They have proven to be an essentialelement that allows estimating the different situations, the various choices as well as their impacts [1].

2. The Business Intelligence Platform

The decision-making information system is a set of data organized in a specific way, easily accessible and appropriate to the decision-making;it is in fact an intelligent representation of these data through specialized tools.

Decision Support Systems, at the level of interaction with the decision maker, allow the refund and the analysis of data from different sources based on the technologies of mass storage, namely the Data Warehousing and OLAP.

Two main functions are intended for the decision-making tools (Figure 1):

1. Collecting and storing, ETL (Extracts, Transforms, Loads) [2], Data Warehousing [3], Datamart, dataweb.
2. Extracting and presenting Data Mining, OLAP.

2.1. Development Tools

The BI platform should provide a set of programmatic development tools and a visual development environment, coupled with a software developer's kit for creating BI applications, in order to integrate them into a business process, and/or embed them in another application. The BI platform should also enable developers to build BI applications without coding by using wizard-like components for a graphical assembly process. [4]

Received January 25, 2016; Revised April 20, 2016; Accepted May 1, 2016

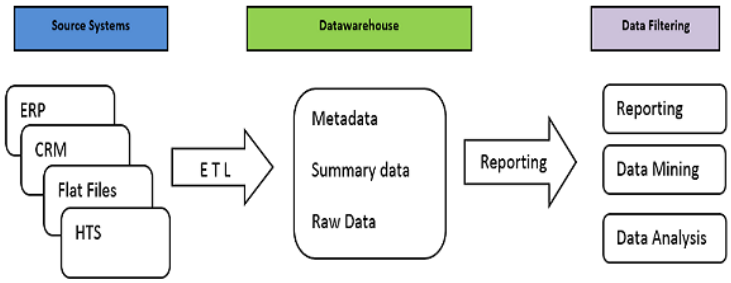

Figure 1. Architecture of the Business Intelligence Platform

2.2. Introduction to Data Mining

The term Data Mining literally means drilling data. As in any drilling, its aim is to be able to extract an element: knowledge. Its concepts are based on the recognition within each company of information hidden in the data repository. It allows, Thanks to a number of specific techniques, it allows making knowledge appear. The Data Mining is the set of techniques and methods aimed at designed for the exploration and analysis of large computer databases in order to detect in these data rules, associations, unknown trends (not fixed a priori), specific structures, thus returning in a concise manner the essential information useful for decision support [5-7].

2.3. The Algorithms of Data Mining

2.3.1. Classification

To overcome the limit of the binary classification in response templates, Charles et al. have applied ada-boost [8] to the Bayesian networks [9] [10] such as algorithms for effective programming to classify customers by their aptitude to respond to bids.

2.3.2. Rules of Associations

The extraction of rules of associations is one of the techniques of popular data mining, which aims to find the relationship between two sets of an object.The names of the developers of algorithms known in this area are Agrawal et al. [11] [12], Zaki Éclat [13], Han et al. [14].

2.3.3. Clustering

The technique of clustering allows toidentify the groups of individuals with similar characteristics. It can be used to distinguish segments of students by sex /institution/region... To achieve this goal several algorithms have been used such as k-means [15], fuzzy c-means [16], the algorithm of Gustafson-Kessel [17] [18] and the algorithm of Gath Geva [19].

2.3.4. Prediction

Predicting the identity of one thing based purely on the description of another related thing itbased on the relationship between a thing that you can know and a thing you need to predict [20]

3. Theoretical Study of Our Application

In this part, we present the class diagram and the sequence for the construction of a data warehouse that has data from the university application known as APOGEE (Application for the organization and management of students and teachers). In effect, the objective of this work isto makean application under the Java programming environment in order to evaluate the algorithms of data mining on data sets that were mined at the level of the data warehouse. The aim is also to present the statistics that describes the data warehouse using graphs and charts.

3.1. The Class Diagram

The class diagram is composed of:

☒***The Student class***: It contains the information of all students enrolled in the institution.

☒***The Institution class***: It contains the name and the acronym of the institution.
☒***The Baccalaureate class***: It contains the type and the series of the student's bachelor degree.
☒***The Path class***: It contains the name, the wording of the chain, the diplomaand the component or thestudent is registered.
☒***The Region class***: It contains the country, the region, and the city to which the student belongs.

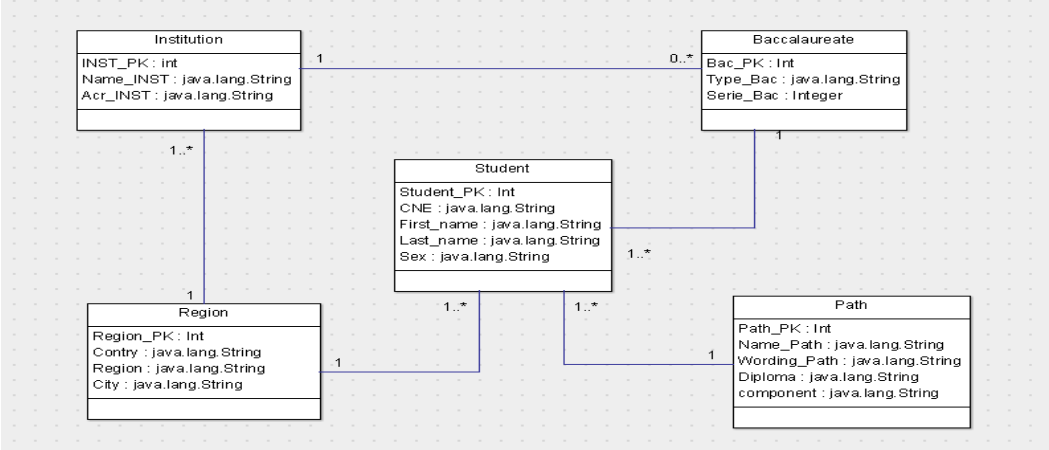

Figure 2. Class diagram of our application

3.2. The Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams are intended to model both computational and organizational processes (i.e. workflows) [21]. Activity diagrams show the overall flow of control. In this section, we will present the Activity Diagram.

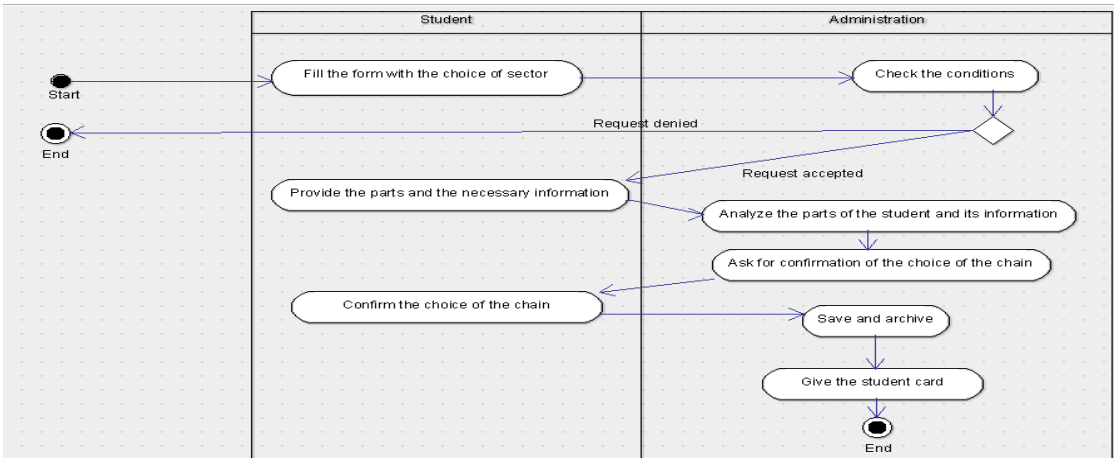

Figure 3. The Activity Diagram

The scenario of our case is the following: The student fills out the enrollment form, the tutor checks and analyzes that all of the information is correct, and requests the student to confirm the choice of the chain. The student confirms the choice of the chain. The tutor records and archives the information of the student and gives the student card.

3.3. Applications of data mining

3.3.1. Classification

Classification uses the decision trees, which offer an outputthat is clear and easy to interpret.

- ✓ Classification of students by age.
- ✓ Classification of students by sex: This criterion is used to determine the sex of student to make clear the gender division and its different rates.
- ✓ Classification of students by sector.
- ✓ Classification of students by note.
- ✓ Classification of students by region.

3.3.2. Algorithm used Clustering K-means

The algorithm of k-means, presented by McQueen in 1967, is an algorithm for partitioning of data falling within the statistics and of the automatic programming (more precisely of the programming which is not supervised). It is a method whose purpose is to divide the observations into K partitions (clusters) in which each observation belongs to the partition with the closest average [22].

3.3.3. Representation of the Application

The application is divided into three parts:

- Homepage
- Data mining part
- Statistics part

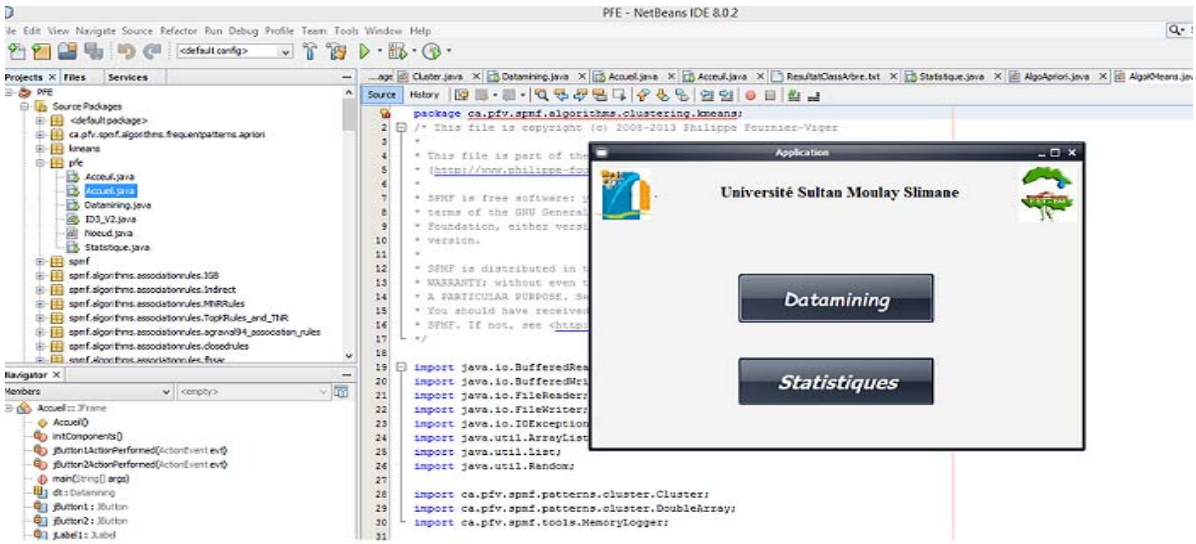

Figure 4. Interface of the Application

3.3.3.1. Data Mining Part

InData Mingpart, three algorithms were used: The classification, the clustering and the rules of associations.

The majority of the work was concentrated on the classification of students by their results; there are two statuses of the results: valid and invalid.

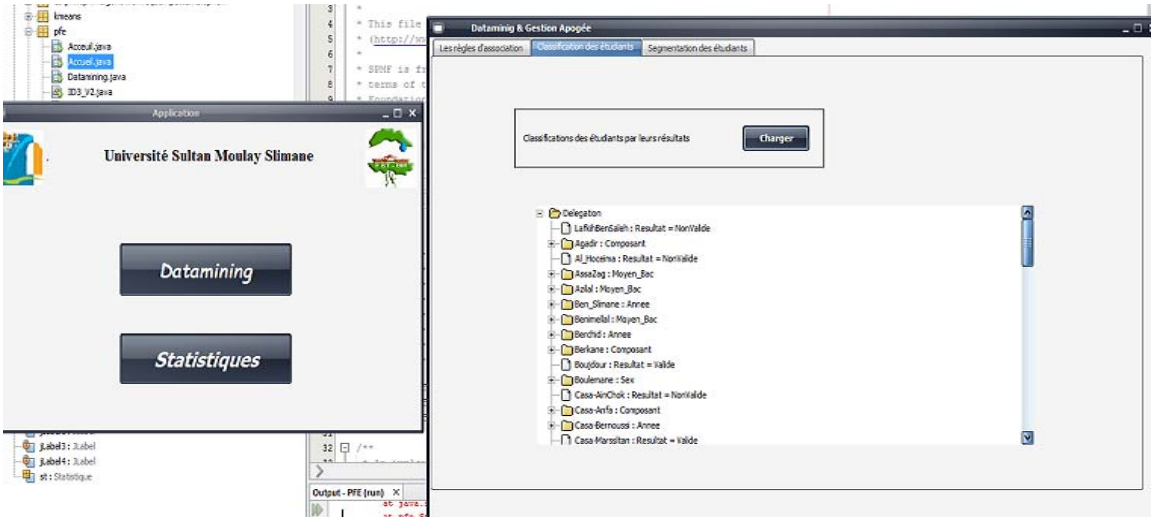

Figure 5. Interface for Data Mining Application

The indicators used in this classification are the region, the study discipline, the sex and the type of baccalaureate.

For example, the contributors who belong tothe regionofTadla and who have a bachelor's degree with distinction have the equivalent of valid module 2.

3.3.3.2. Statistic Part

The statistical part is divided into four topics:

- **Inscription**: displaying students enrolled by their regions and Baccalaureate distinction
- **Validation**: displaying students who validated by their institutions and by years of graduation
- **Sex**: display of students enrolled by their sexes
- **Bourse**: display of students' scholarship according to their regions

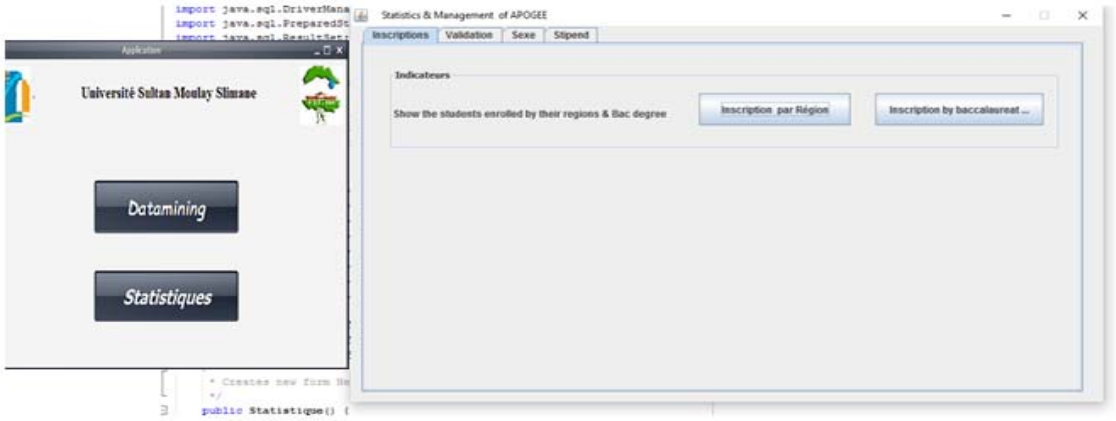

Figure 5. Interface for statistic application

a. The Inscription Part

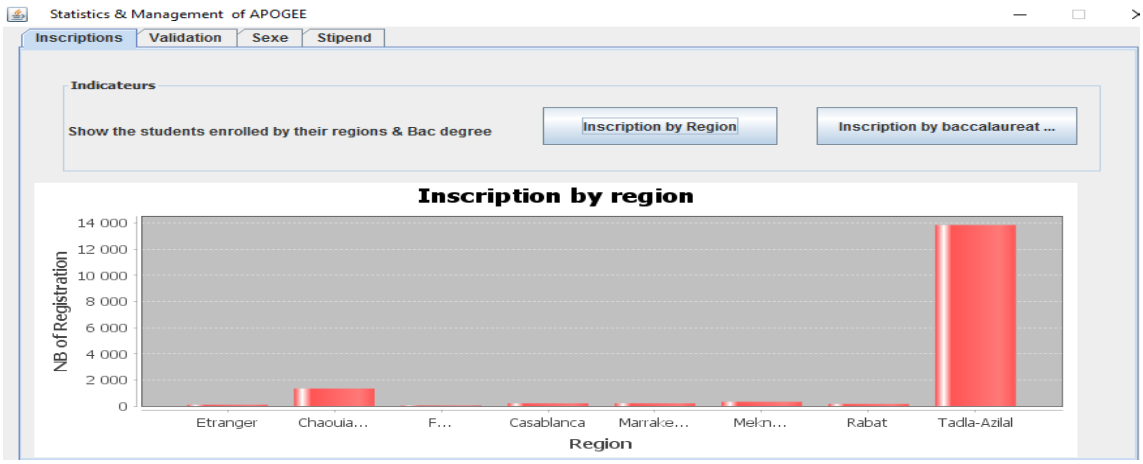

Figure 6. Registration of Students by Regions

The majority of students come from the regions of Tadla-Azilal and Chaouia.

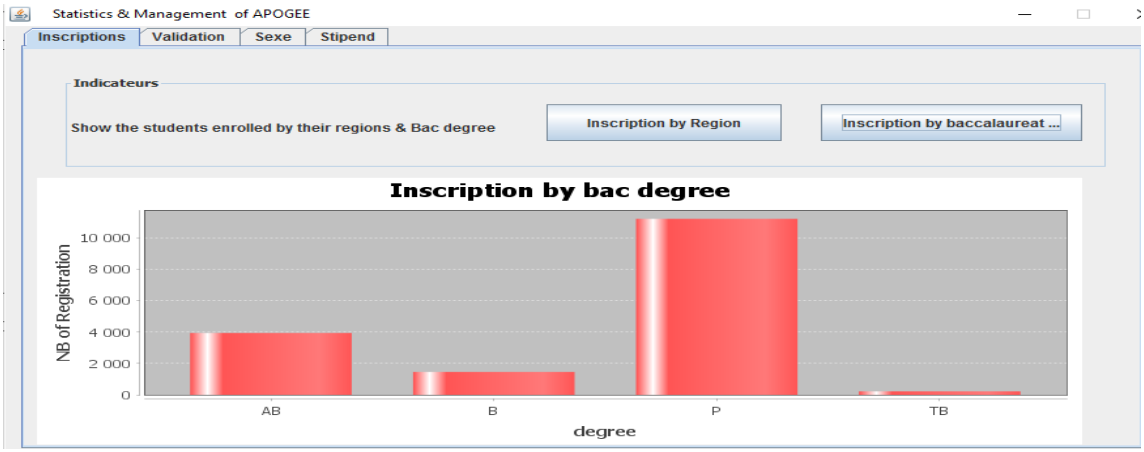

Figure 7. Registration of Students by Bac

This part includes four types of honors: TB: Very Good (Ist class)B: Good(upper IInd class) AB: good enough (Lower IInd class) P: passable (IIIrd class)

The majority of students have the degree: very good und good enough

b. The validation part

The University Sultan Moulay Slimane is comprised of the following institutions:

FLSH: Faculty of Letters and Human Sciences
FST: Faculty of Scienceand Techniques
FP : Multidisciplinary College
EST: Higher College of Technology

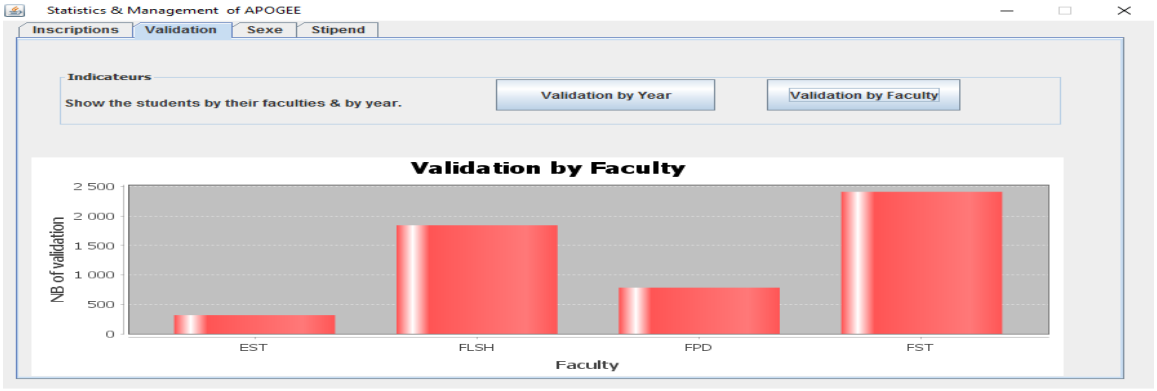

Figure 8. validation of Students by Faculty

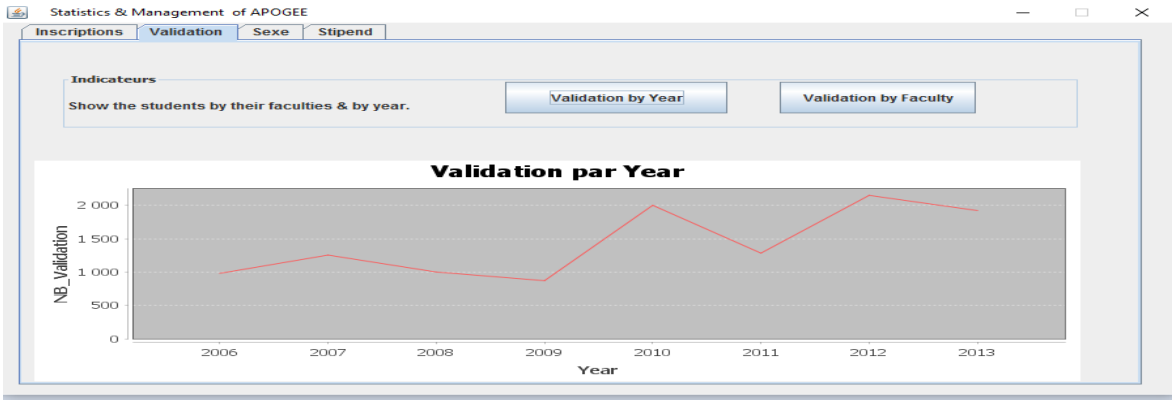

Figure 9. validation of Students by year

The Figures 8 and 9 illustrate the number of students who have validated the exams by Faculty and by year.

c. Sex of students part

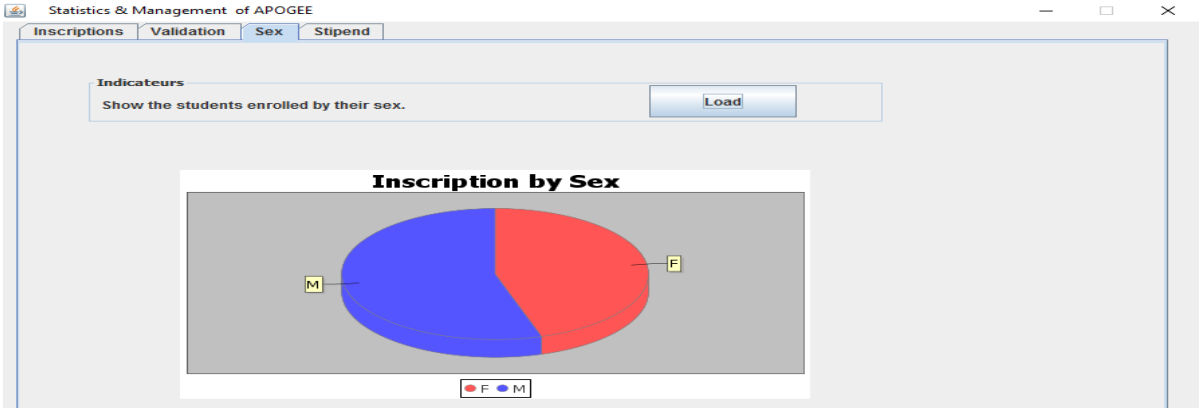

Figure 10. Registration of Students by sex

d. The students fellows part

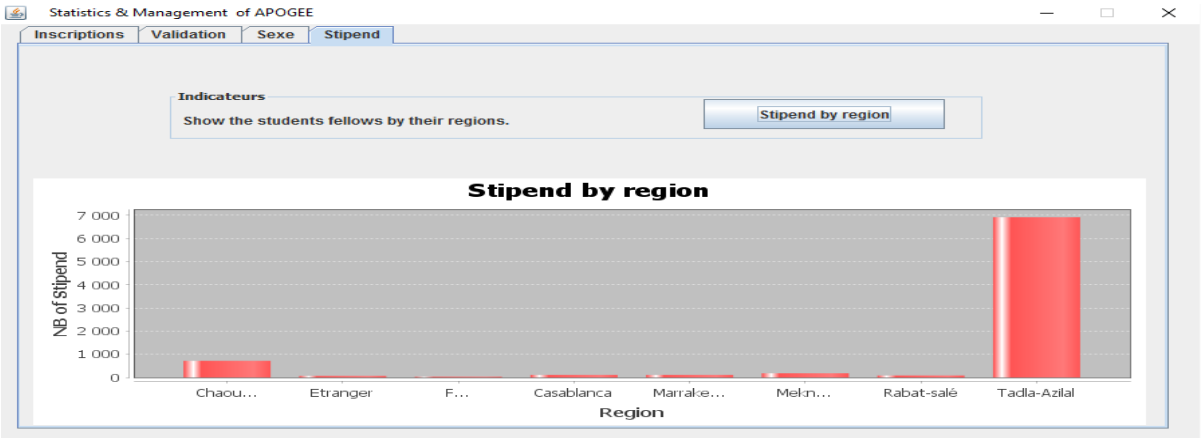

Figure 11. The Students fellows by their regions

4. Conclusion

This article is intended to put in place a Decision Support System for the management of students and shows how it is possible to exploit the existing databases in order to design a data warehouse that integrates the design and the needs of end-users.

Actors-oriented trades' databases allow giving suggestions in decision-making. Finally, we have designed an application under the Java programming environment in order to evaluate the algorithms of data mining on the basis of datasets that were extracted from the data warehouse. The article also presents the statistics that describes the data warehouse using graphs and charts.

References

[1] Lévine P, Pomerol J. Editors. *Systèmes interactifs d'aide à la décision et systèmes experts*. Hermès. 1989
[2] Golfarelli M, Rizzi S. *Editors. Data Warehouse Design, Modern Principles and Methodologies*. McGraw-Hill Education. 2009.
[3] Inmon W.H. Editors. Building the data warehouse. Wiley. 1996.
[4] Wieder B, Ossimitz M. *The impact of Business Intelligence on the quality of decision-making – a mediation Model.* Conference on ENTERprise Information Systems. CENTERIS. 2015; 64: 1163–1171.
[5] Tuffery S. Editors. Datamining et statistique décisionnelle : l'intelligence des données.Editions technip. 2012.
[6] Roy B, Bouyssou D. Editors. Aide multicritère à la décision : *méthodes et cas. Economica*.1993.
[7] Siddartha .B.*Direct Marketing Response Models using Genetic Algorithms*. American Association for Artificial Intelligence.1998: 144-148.
[8] Freund Y, Schapire RE. *Experiments with a new boosting algorithm*. In Proceedings of the Thirteenth International Conference on Machine Learning.1996: 148 – 156.
[9] Elkan C.*Boosting and naive Bayesian learning*. Department of Computer Science and Engineering University of California. 1997.
[10] Langley P, Iba W, Thompson K. *An Analysis of Bayesian Classifiers*. In Proceedings of the Tenth National Conference on Artificial Intelligence. Moffet Field. 1992: 223 – 228.
[11] Agrawal R, Imielinski T, Swami A. *Mining association rules between sets of items in large databases*. In Proceedings of the ACM SIGMOD Conference. Washington.1993: 207-216.
[12] Agrawal R, Srikant R .*Fast Algorithms for Mining Association Rules*. In Proceedings of the 20th VLDB conference.Santiago.1994: 487-499.
[13] Zaki MJ. Scalable algorithms for association mining. *IEEE Transactions on Knowledge and Data Engineering*. 2000; 12(3): 372–390.

[14] Han J, Pei J, Yin Y. *Mining Frequent Patterns without Candidate Generation*.SIGMOD Proceedings of The ACM SIGMOD international conference on Management of data. New York. 2000; 29(2): 1-12.
[15] Kaufman L, Rousseeuw P J, Dodge Y. Statistical Data Analysis Based on the L1–Norm and Related Methods.Elsevier Science.1987: 405–416.
[16] Dunn JC. *A* Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics.*1973; 3(3): 32-57.
[17] Gustafson D E., Kessel W C. *Fuzzy Clustering with a Fuzzy Covariance Matrix*, Decision and Control Including the 17th Symposium on Adaptive Processes, IEEE Conference. San Diego. 1978; 17:761-766.
[18] Babuka R, Van der Veen PJ, Kaymak U. *Improved covariance estimation for Gustafson-Kessel Clustering*. IEEE Conference. Honolulu. 2002; 2: 1081-1085.
[19] Sarwar B†,Karypis G‡, Konstan J†, Riedl J†.*Recommender Systems for Large-scale E-Commerce: Scalable Neighborhood Formation Using Clustering*.The International Conference on Computer and Information Technology. Dhaka. 2002.
[20] Zhang H, Zhao G, Chen L, Liu B. Short-Term Prediction of Wind Power Based on an Improved PSO Neural Network.*TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2014 ; 12(7): 4973-4980.
[21] Rumbaugh J, Jacobson I, Booch G.Editors. The Unified Modeling Language Reference Manual. Addison-Wesley. 1999.
[22] Sai Chandana B, Srinivas K, Kiran Kumar R. Clustering Algorithm Combined with Hill Climbing for Classification of Remote Sensing Image. *International Journal of Electrical and Computer Engineering (IJECE)*. 2014; 4(6): 923-930.