EMPLOYEE DATA MINING: ANALYSIS OF DATA SETS FOR IMPROVING HUMAN RESOURCE PRACTICES

[1]JULIET V. MENOR, [2]JULIE D. ROSAL

Pangasinan State University, San Carlos City Campus, San Carlos City, Pangasinan, Philippines
E-mail: [1]psujulietmenor@gmail.com, [2]andjuliedrosal@gmail.com

Abstract - Human resources is used to describe both the people who work for a company or organization and the department responsible for managing resources related to employees .It is very important for the company's to determine the rate of attrition in order for HR Manager to plan and prepare the required number of Human Resources. The data set used from the University from the year 1980 to April 2017 was analysed to find out the summary of employee's reason for human resources practices. The main objectives of the study is to determine the rate of employee retention at Pangasinan State University, the study includes age, gender, employment status, appointment status, marital status and length of service. Second, to identify the type of employee attrition by using data mining analysis. The result describes the frequency of attrition status, attrition reasons and other attentiveness of variables. Further, in this paper, a data mining analysis was used to visualize the data from Pangasinan State University.

Keywords - Human Resource, Classification or Clustering, Data Mining Analysis and Visualization

I. INTRODUCTION

Human resources (HR) are one of the most important "assets" of your business. The individuals who make up your organization go a long way toward driving its success, whether they are employees, managers, consultants or others. HR refers to the recruitment and management of personnel, in compliance with the law. Thus, the practice of HR requires knowledge of the laws that affect your workplace - at the federal, state and local levels. Human resources is used to describe both the people who work for a company or organization and the department responsible for managing resources related to employees.

Data Mining is one of AI technology that has been developed for exploration and analysis in large quantities of data to discover meaningful patterns and rules. In actual fact, such data in HR data can provide a rich resource for knowledge discovery and decision support tools. A data mining process aims to improve the probability of predicting a result in comparison with a purely random 'guess'. In such successful models can be compared to other approaches such as human intuition.

The human resource practices or employee attrition denotes to the loss of employees through a number of circumstances, such as resignation layoff and retirement. The attrition reasons may be either voluntary or involuntary, though employer-initiated events such as layoffs are not typically included in the definition. The company has its own standards for reasonable attrition charges, and these charges can also contrast among capable and incapable positions. Due to the costs related with training new employees, any type of employee attrition is typically seen to have a monetary cost. It is also possible for a company to use employee attrition to its advantage in some circumstances, depend on it's to control labor charges without delivering mass layoffs. [1]

There are many different ways for a company to lose employees, most of which are typically taken i account to ensure that the organization is able to operate efficiently. Attrition refers to the loss of employees due to reasons other than firing and other employer-initiated events.[2] This means that an employer has no direct control over how many personnel are lost to employee attrition. Retirement is one major cause of employee attrition, and since people tend to retire around a specific age this is a factor that can be accounted and planned for. Other causes of employee attrition, such as personnel who quit due to prolonged illness, dissatisfaction with the company, or other reasons, can be more difficult to estimate.

It is very important for the company's to determine the rate of attrition in order for HR Manager to plan and prepare the required number of Human Resources. The data used in this research is taken from the University [3].

The main objectives of the study is to determine the rate of employee retention from University in terms of age, gender, employment status, appointment status, marital status and length of service. Second, to identify the attrition type of employee by using data mining analysis.

III. RELATED WORKS

Attrition is the firm reduction in employees of an organization through method of retirement, resignation or death (Elfers et al., 2006; and Goswani Jah, 2012) [4]. In other words, attrition could be the number of employees quitting their jobs which includes both voluntary and involuntary separation.

The problem of attrition is insurmountable whether in the education sector or any other profession, some attrition is inevitable. This is because, some teachers will retire, some others will quit for reasons that is personal, and some others will be dismissed from their jobs whilst some others are encouraged to leave the profession. Henke et al. (2001) found that about 50 percent of those engaged as classroom instructors quit within a five years period, and the best and brightest teachers are often the first to leave [5]. This happening has prompted the concern of why teachers leave their jobs within few years of entering the profession. McKee (2003) argues that employee attrition is an important challenge that cannot be underestimated because it results in reduction of personnel in schools [6]. This is more evident in the fact that almost half of all new teachers quit the job in less than five years of starting. This creates a difficult situation where more teachers quit compared to those coming into the profession.

However, some of the challenges of employee attrition may be linked to career progression of the people involved (Barmby, 2006; Borman and Dewling, 2008) [7]. In this case, attrition could not be said to be bad. Nevertheless, majority of the reasons that result to employee attrition is viewed from variables such as poor compensation, lack of management encouragement, not too conducive work atmosphere, issues connected to co-employee and collegiality with peers. There is need to understand thoroughly the factors that contribute to employee attrition so as to be able to provide a conducive work environment for them. Empirical literature has identified many causes of employee attrition in recent years. Guarino et al. (2006) stresses that some of the studies revealed that the problem is due to standard of human capital accumulated by employees [8]. Some evidence show that employees who acquire higher certificates obtained in more lucrative courses and in prime universities and colleges would most likely not stay too long in the teaching profession and administrative work. But, for degrees that are not of high market value, the probability of such teaching staying is higher. Another group of researchers show that attrition can be subcategorized into many kinds, each affected variously by human and social capital (Azuka, 2003; Long et al., 2012a) [9]. Some educators quit fulltime classroom teaching either for professions in an unconnected areas if they quit the teaching profession entirely.

Ingersoll (2003) opines that other individuals can remain working in the profession but move to other schools. Another view suggests that some may temporarily quit the profession perhaps to have children, and then come back to their job afterwards [10]. Anderson and Olsen (2005) found that retirement is naturally perceived among the primary reasons for teacher attrition. However, the amount of teachers quitting through this channel show that it may not be a course for concern [11]. For example, Ingersoll (2003) investigates teacher turnover and cites that retirement of teachers is not an essentially domineering reasons for attrition than job dissatisfaction or to leaving to follow another career. Hence, the literature indicates that a substantial amount of teachers resign their current job since there does not seem to be hope for change. Moreover, Job satisfaction or dissatisfaction is always mentioned and emphasized as essential on studies relating to teacher attrition and retention (Stockard and Lehman, 2004; Ingersoll, 2003; Long et al., 2012b) [12]. In order to find out the common problems identified for teacher turnover and retention rate among beginning teachers, Stockard and Lehman (2004) found the following as the reasons and causes; dissatisfaction of workers with current job, unsatisfactory conditions of work and inadequate remuneration, not enough support from family members, management challenges, colleagues and the public, discipline, challenges of attendance, rising size of classes, inadequate encouragement of students and inadequate opportunities for teachers to partake in important decisions that affects their schools.

Boyd et al. (2007) posits that many factors can contribute to employee attrition. The researcher mentions further that among reasons that determines whether a teacher will remain on his job or quit the profession depend on individual factors, but these may vary from one company to another [13]. Bobbitt et al. (1994) posits that not minding these, there are factors that are peculiar to different areas and countries, the main cause of teacher intention to leave is the low salaries and unsatisfactory working conditions [14]. Reports in sub-Saharan Africa suggests that attrition and turnover is associated with the HIV and AIDS epidemic and also due to such reasons as lack of adequate salaries, allowances, housing and promotion (Dinham, 1994; Boyd et al., 2007) [15]. Attrition and retention of classroom educational managers is a complicated issues, with reasons to remain or leave usually linked to personal variables, still, some themes and patterns may be connected (Boyd et al., 2006) [16]. A variety of constructs can be related to the challenges of teacher attrition, nonetheless, the variables may vary from one country to the other. The dimensions in USA may not be the same in Malaysia or in Nigeria, but not withstanding the differences in these variables that are common to all countries as reported in past studies.

Long et al. (2012a) opines that turnover of employees has an enduring impact on the productivity of organizations [17]. The authors mentioned further that when they quit, there would be an urgent need for replacement which results to the organization incurring certain amount of tangible expenditure

(Azuka, 2003; Baike, 2002; Achor et al., 2009; Long et al., 2012b) [18]. However, such effects could range from expenses that would be incurred to replace such employee and the problem of managing culture shock of the new employees.

III. METHODOLOGY

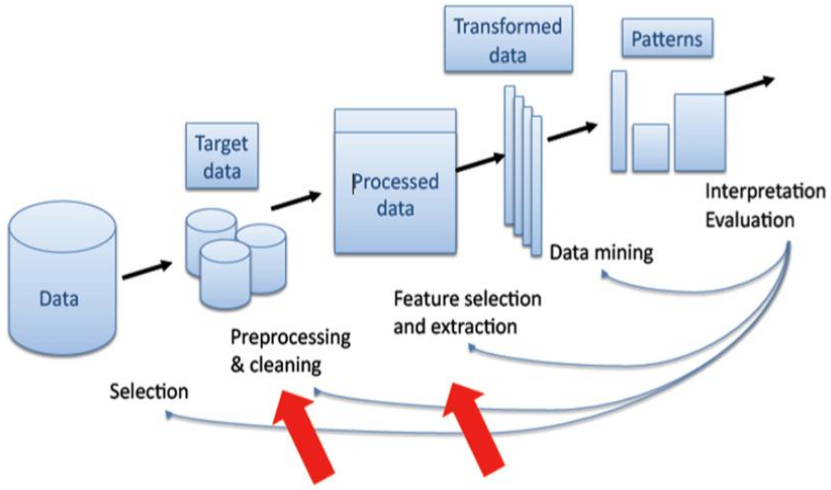


Figure 1. Data Mining Model

The basic model of employee attrition works by clustering or classifying employee profiles based on various attributes such as age, gender, employment status, marital status, appointment status, department and termination reasons can generates various levels of attrition. Sometimes, other parameters like length of service over the years stayed are also taken into consideration**.**

The basic model of employee attrition works by clustering or classifying employee profiles based on various attributes such as age, gender, employment status, marital status, appointment status, department and termination reasons can generates various levels of attrition. Sometimes, other parameters like length of service over the years stayed are also taken into consideration.

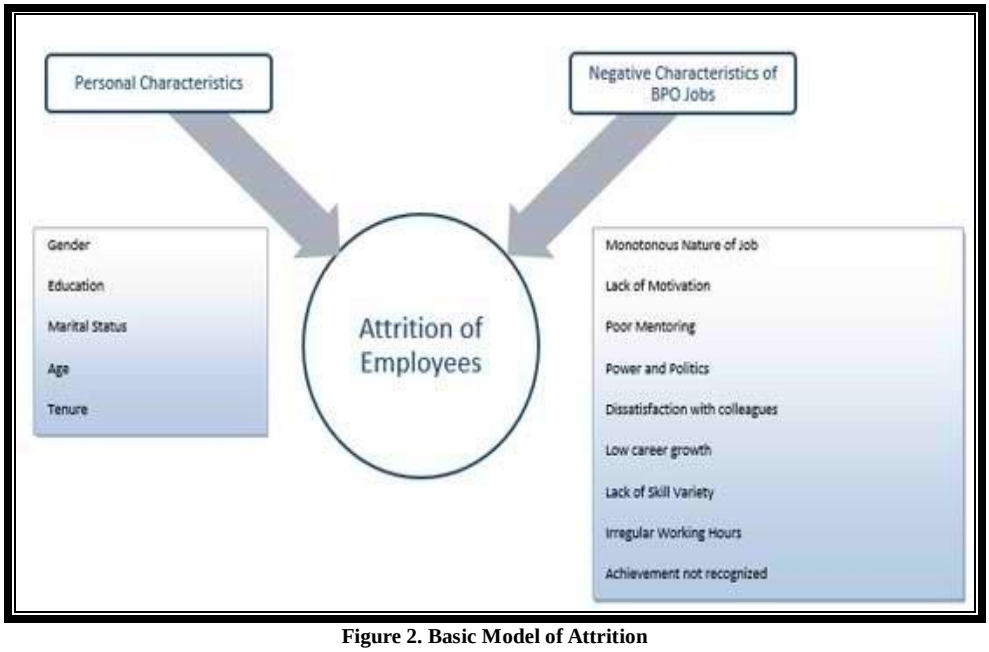


Figure 2. Basic Model of Attrition

This study used the descriptive type of research. The descriptive research as defined by Walliman (2011) contains the data, examination, analysis and visualization. The descriptive research a method which deals to describe the factor of attrition on the profiles of the respondents through attrition reasons using data mining

techniques in terms of age, gender, employment status, appointment status, marital status and length of service [19].

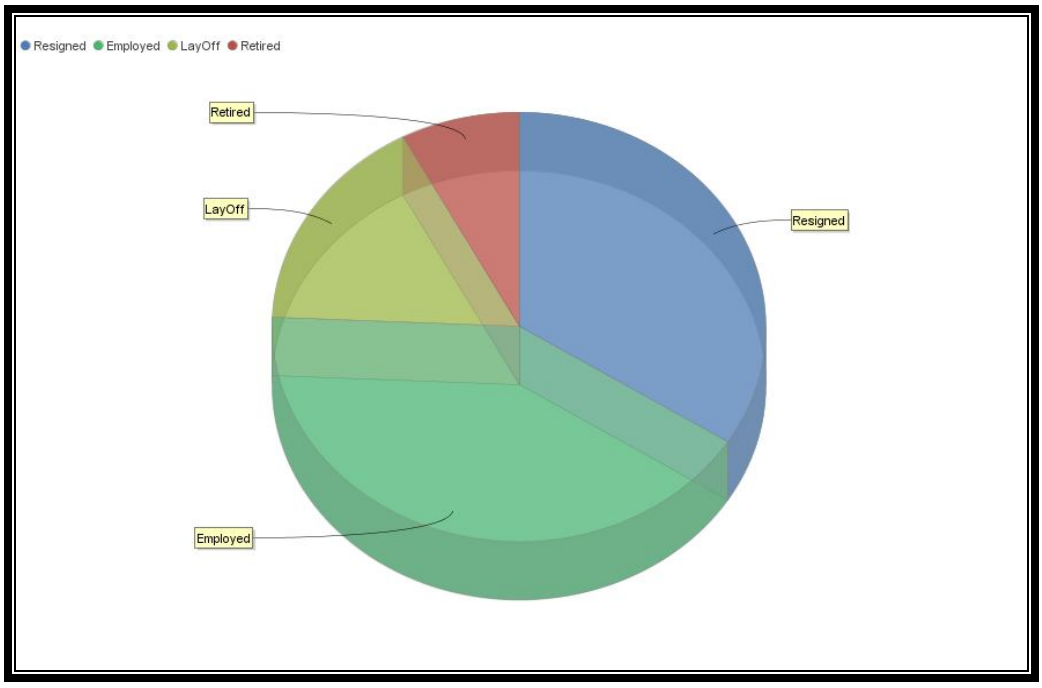
IV. RESULTS AND DISCUSSIONS

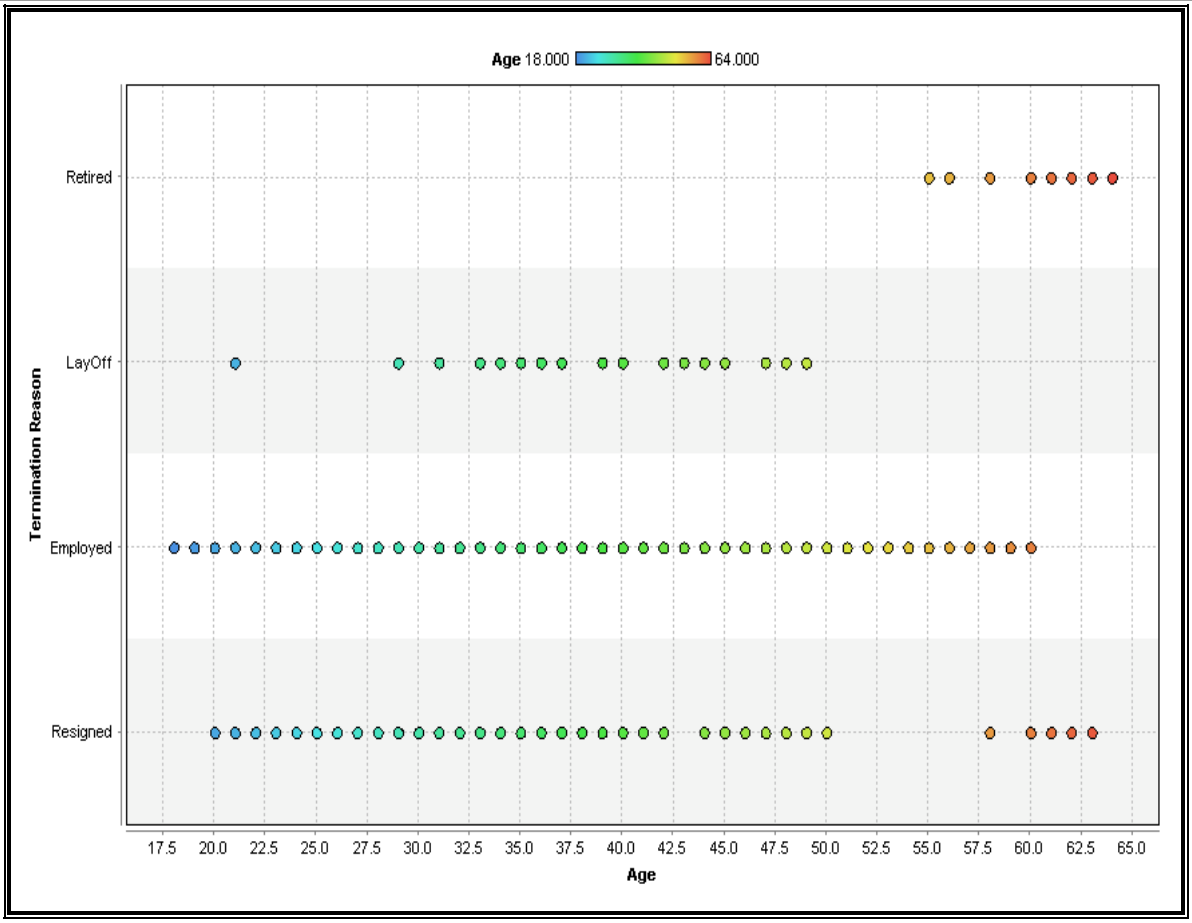
This section discusses the data mining results based on the different variables. The raw data from select University in the Philippines was analysed to find out the attrition of employee's reason. The result describes the frequency rate of attrition status, attrition reasons and other attentiveness of variables.

Age	Integer	0	Min 18	Max 64	Average 37.235
Employment Status	Polynominal	0	Least Terminated (237)	Most Active (1170)	Values Active (1170), Terminated (237)
Termination Reason	Polynominal	0	Least Retired (24)	Most Employed (1170)	Values Employed (1170), Resigned (186), ...[2 more]
Department	Polynominal	0	Least ADMIN (483)	Most ACAD (924)	Values ACAD (924), ADMIN (483)
Status Year	Polynominal	0	Least 9/29/2016 (1)	Most 01/07/2016 (205)	Values 01/07/2016 (205), 03/03/2017 (146), ...[612 more]
Length of Service	Integer	0	Min 0	Max 37	Average 6.810
POSITION	Polynominal	0	Least VP.ACAD (1)	Most INSTRUCTOR I (494)	Values INSTRUCTOR I (494), ADMAIDE I (111), ...[99 more]
Education	Integer	0	Min 1	Max 5	Average 2.912
Education Field	Polynominal	0	Least Math (10)	Most Business (640)	Values Business (640), Marketing (211), ...[8 more]
Appointment Status	Polynominal	0	Least Temporary (104)	Most Contractual (547)	Values Contractual (547), Regular (515), ...[2 more]
MaritalStatus	Polynominal	0	Least Separated (1)	Most Married (743)	Values Married (743), Single (587), ...[4 more]
Gender	Polynominal	0	Least Female (619)	Most Male (788)	Values Male (788), Female (619)
TrainingTimesLastYear	Integer	0	Min 0	Max 6	Average 2.796
YearsAtUniversity	Integer	0	Min 0	Max 37	Average 6.180
YearsInCurrentRole	Integer	0	Min 0	Max 17	Average 3.943
YearsSinceLastPromotion	Integer	0	Min 0	Max 15	Average 2.107

Table 1. Profile of Employee Attrition

Based on the Table 1 shows the employee profile that run and analysed and visualized the dataset using rapid miner and orange projection software.





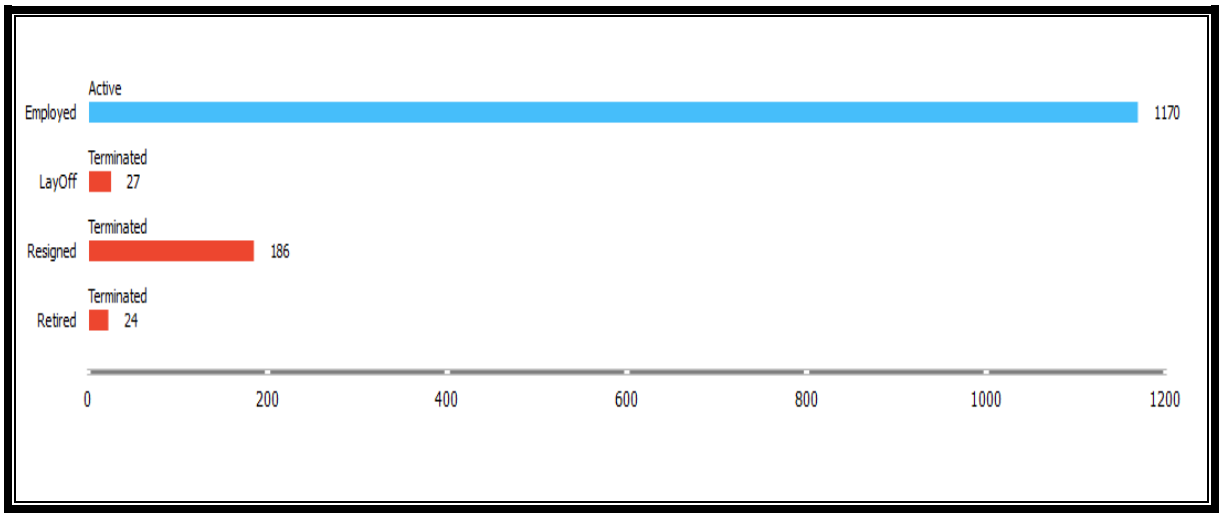


Figure 1. Employees Termination Reasons

It can be observed from figure 1 the status of the employees. It can be noted from the graph that majority are employed. This followed by 1170 frequency rate while some because of employment lay off comprising about 27 frequency rate, resigned comprising about 186 frequency rate and lastly due to retirement which is represented by 24 frequency rate.

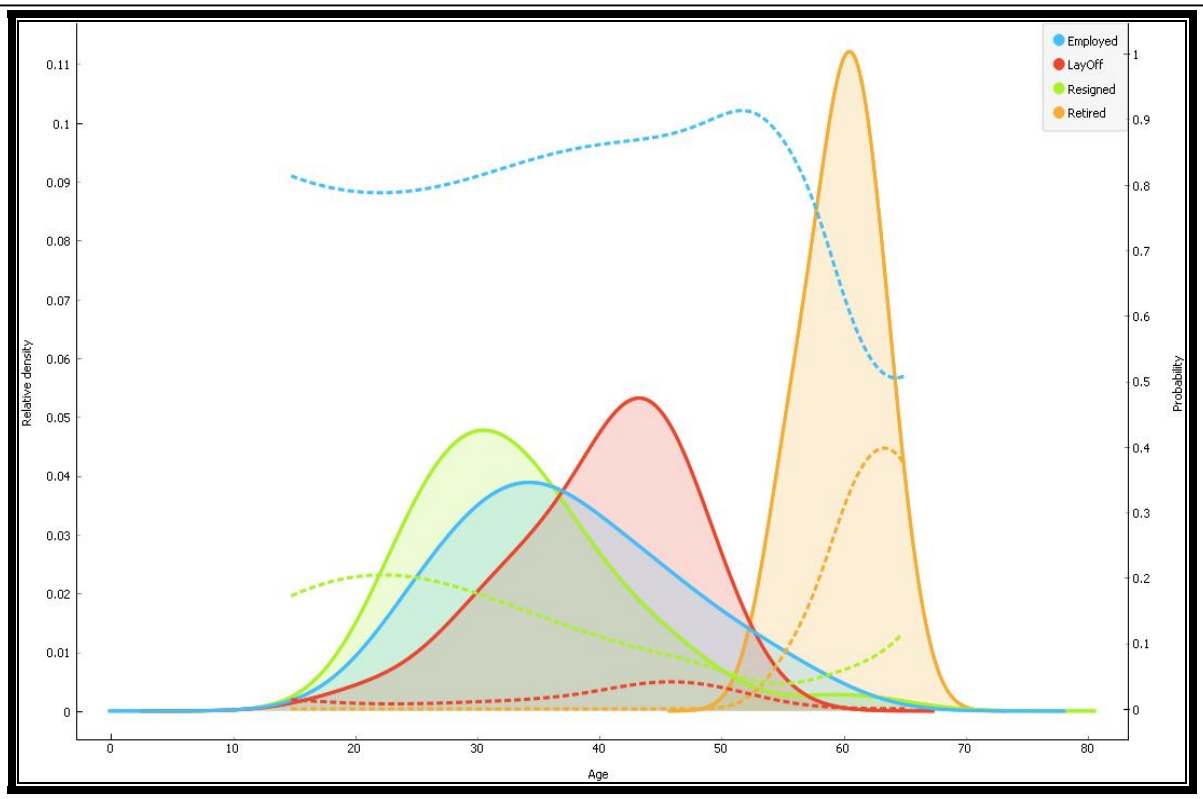


Figure 2. Age with most Employee Attrition

As observed from Figure 2 it can be noted that the age with the most number of attrition is 37 years, subsequently this is followed by 18 years, based on the analysis of the data, age range from 25 to 37 years old are the majority who leave their job.

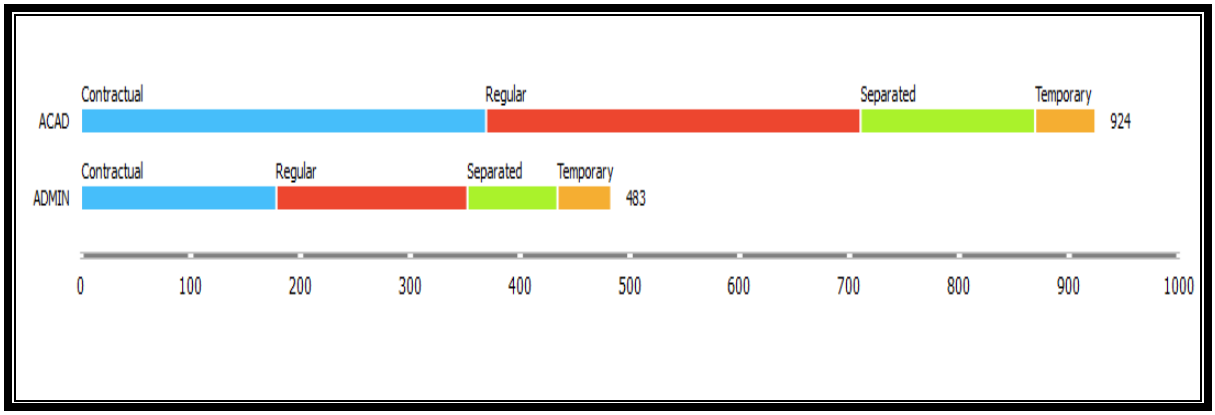


Figure 3. Department with most Employee Attrition

The University comprises two department, namely the Academic and Administrative department. Based on the analysis of the data, figure 3 showed that majority leave their job because of attrition reasons such as contractual, regular, and temporary and separated which comprises about 924 frequency rate, while some from Admin department comprising about 483 frequency rate.

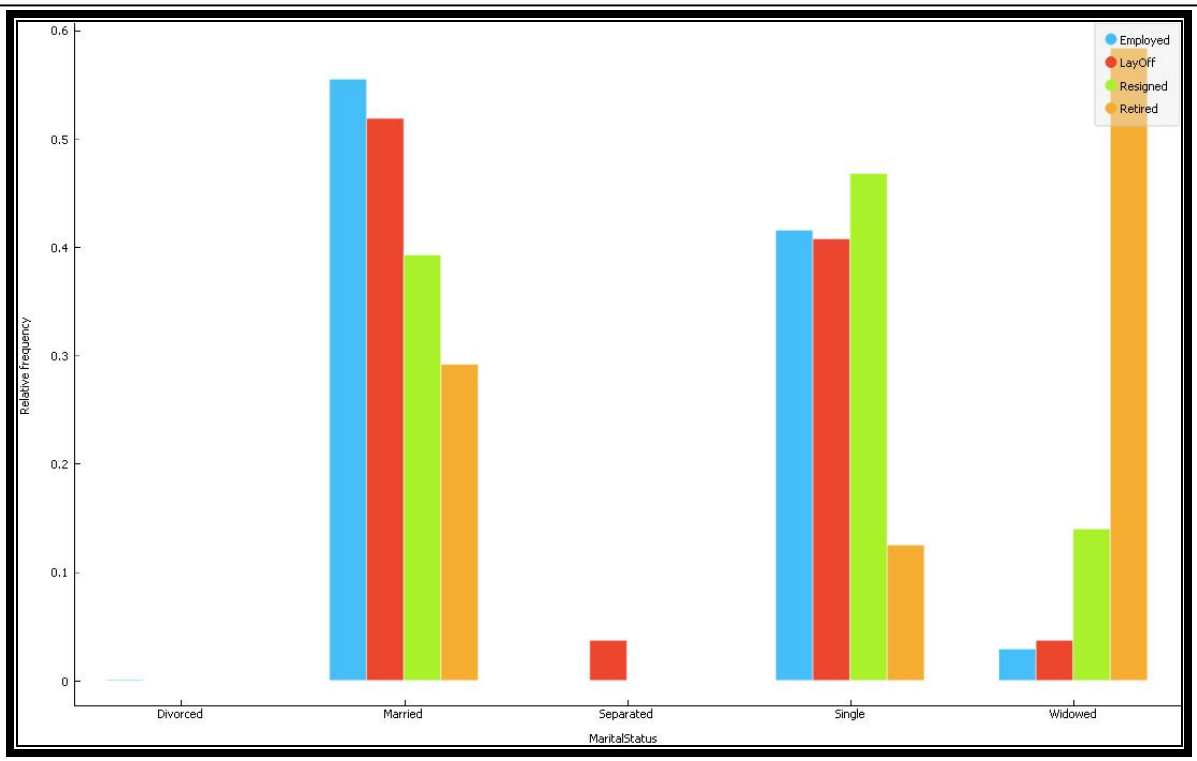


Figure 4. Marital Status of Employee Attrition

The employee marital status comprises the single, married, separated, divorced and widowed. Based on the analysis of the data, figure 4 showed that there are 649 single, 587 married, 34 widowed, 1 divorced, and no separated employee.

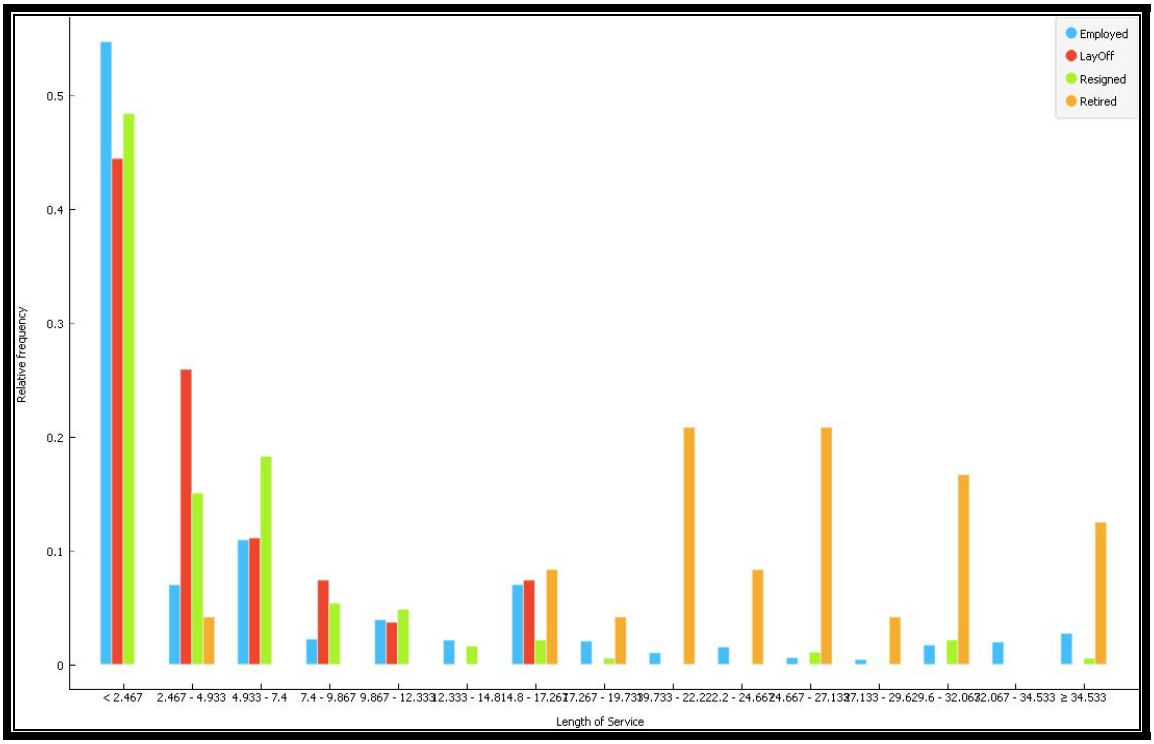


Figure 5. Length of Service

In term of length of service Figure 5, shows that some employees leave their job in a few month time while some reaches 37 years of service.

This study estimate and visualize the general attrition recovered from the selected termination reasons. Each reason estimated attrition is represented by a circle mapped by emotions. The termination reasons is plotted in

red circles while active employees are mapped in blue circles. The stimulation is represented as red circles which indicate that the brighter the circle, the more active are the terminated.

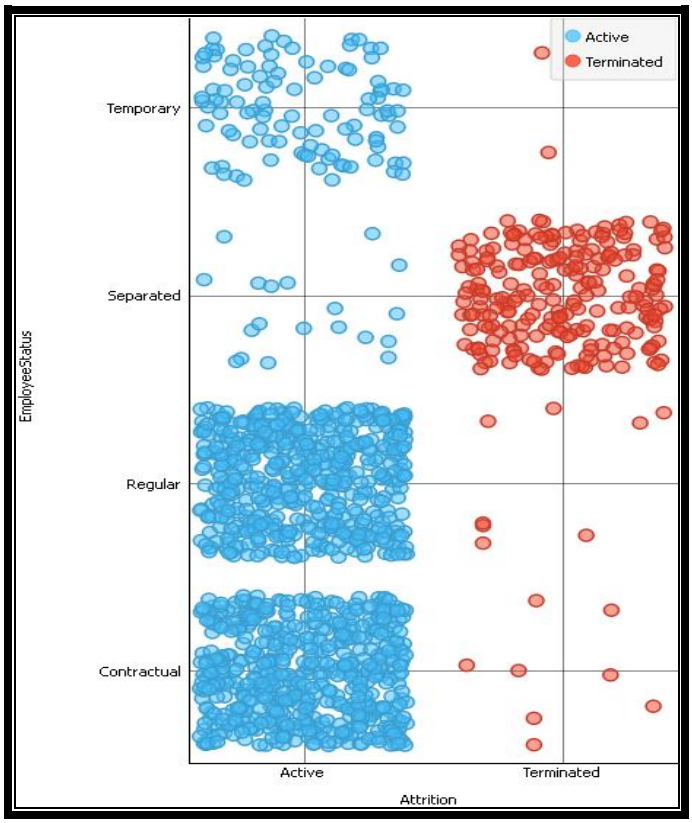


Figure 6. Employee Attrition

Figure 6 represented the Employee Attrition of the selected variables. The graph shows the general employees attrition visualized in scatter plot diagram with desire and stimulation represented in horizontal and vertical axis. The 2-dimensional distribution summarizes the overall status of employee's attrition.

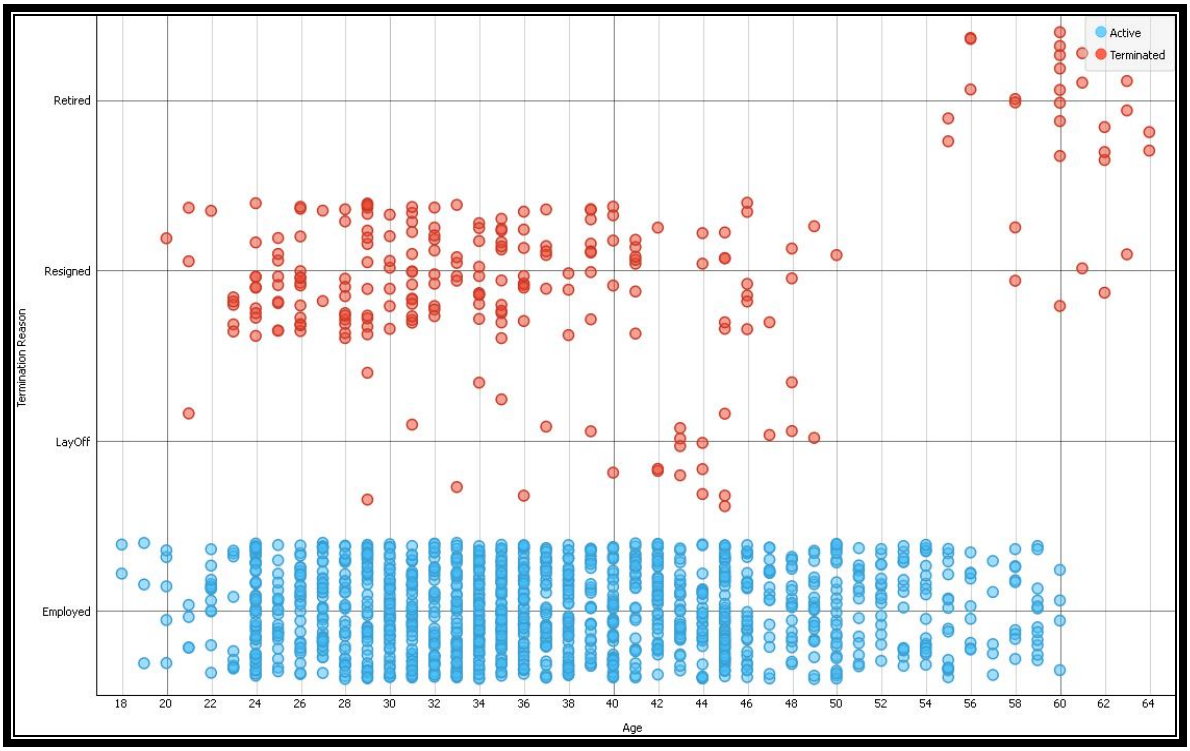


Figure 7. Age of Employees

Figure 7 shows the age attrition of the selected variables. The graph shows the general age attrition visualized in scatter plot. The distribution summarizes the overall age attrition that most age range from 25 to 40 years of age who leave.

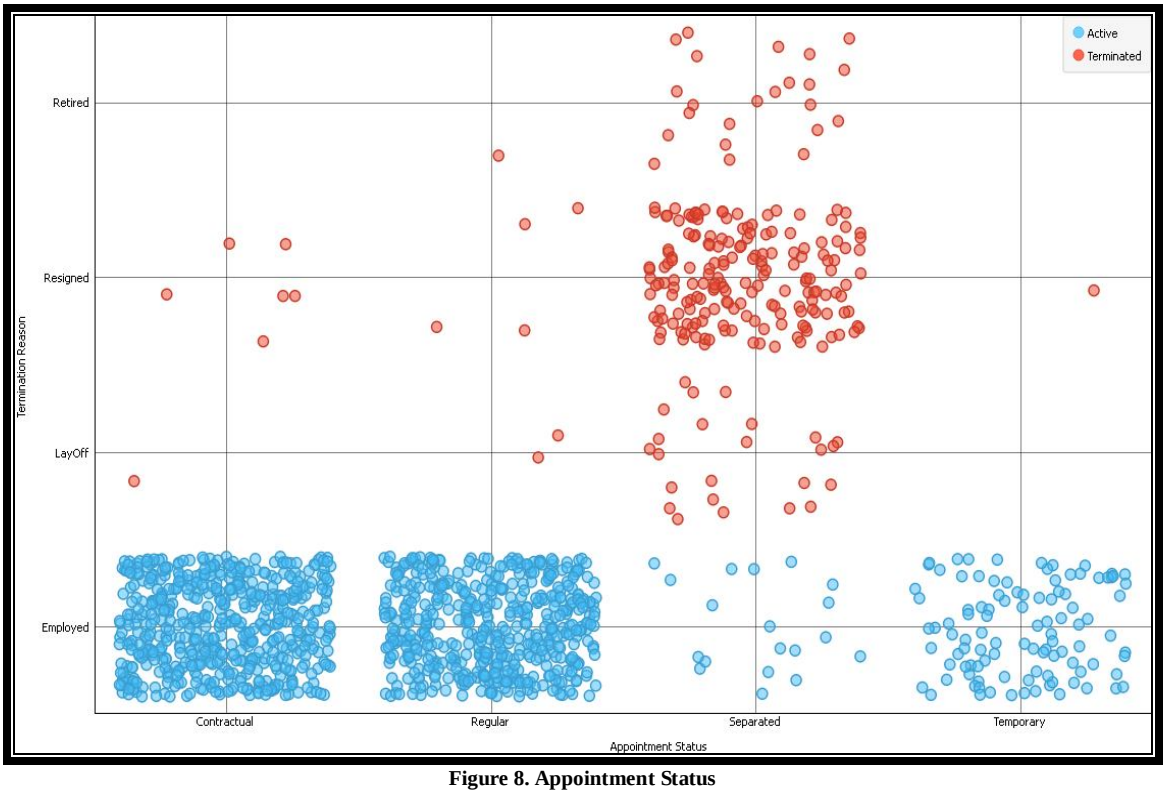


Figure 8. Appointment Status

Figure 8 shows the appointment status common with the most attrition is single, based on the visualization of the data analysis.

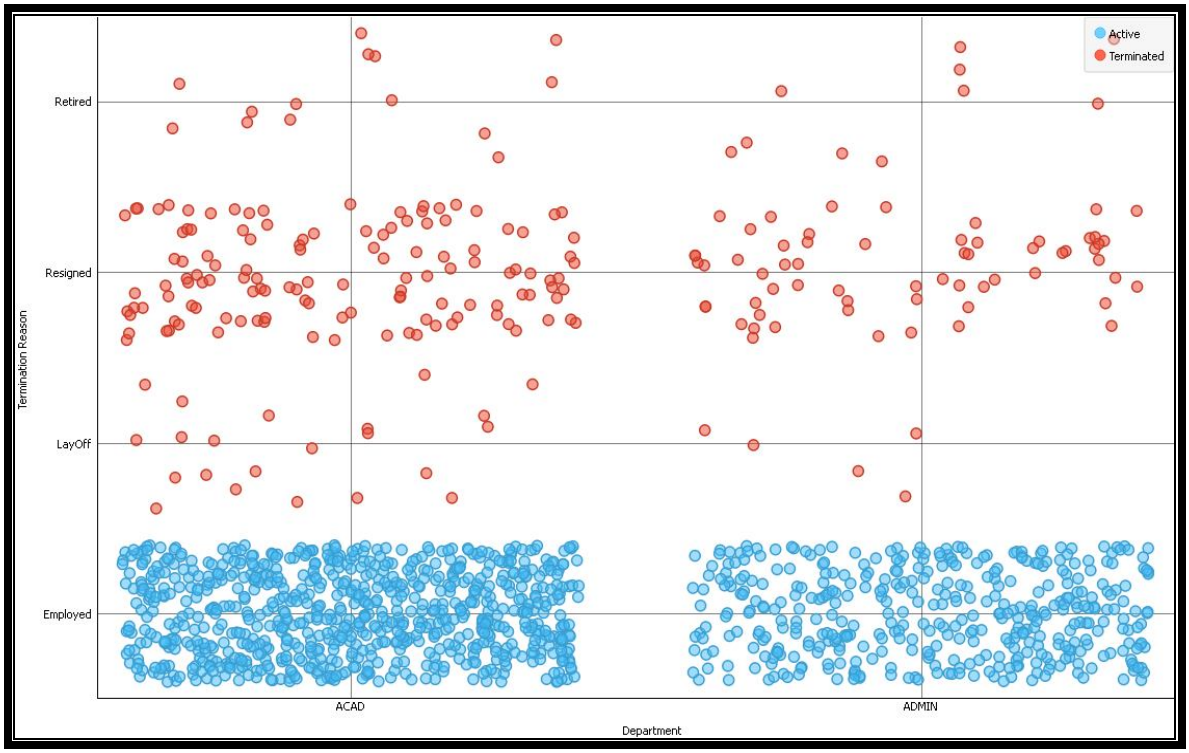


Figure 9. Department

Figure 9 shows the department with the greatest attrition is the academic department which composes of the faculty members from the different colleges of the university.

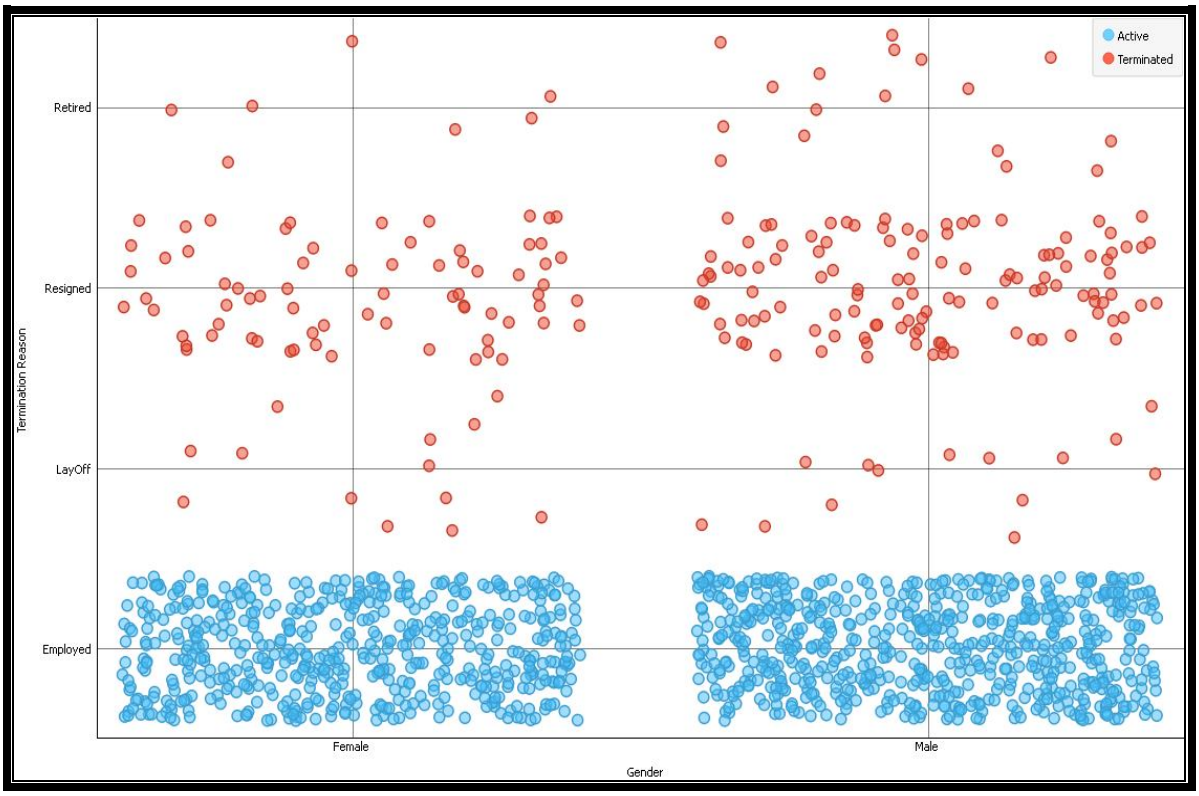


Figure 10. Gender

Figure 10 shows the gender with an highest attrition is male than female, based on the visualization of data analysis.

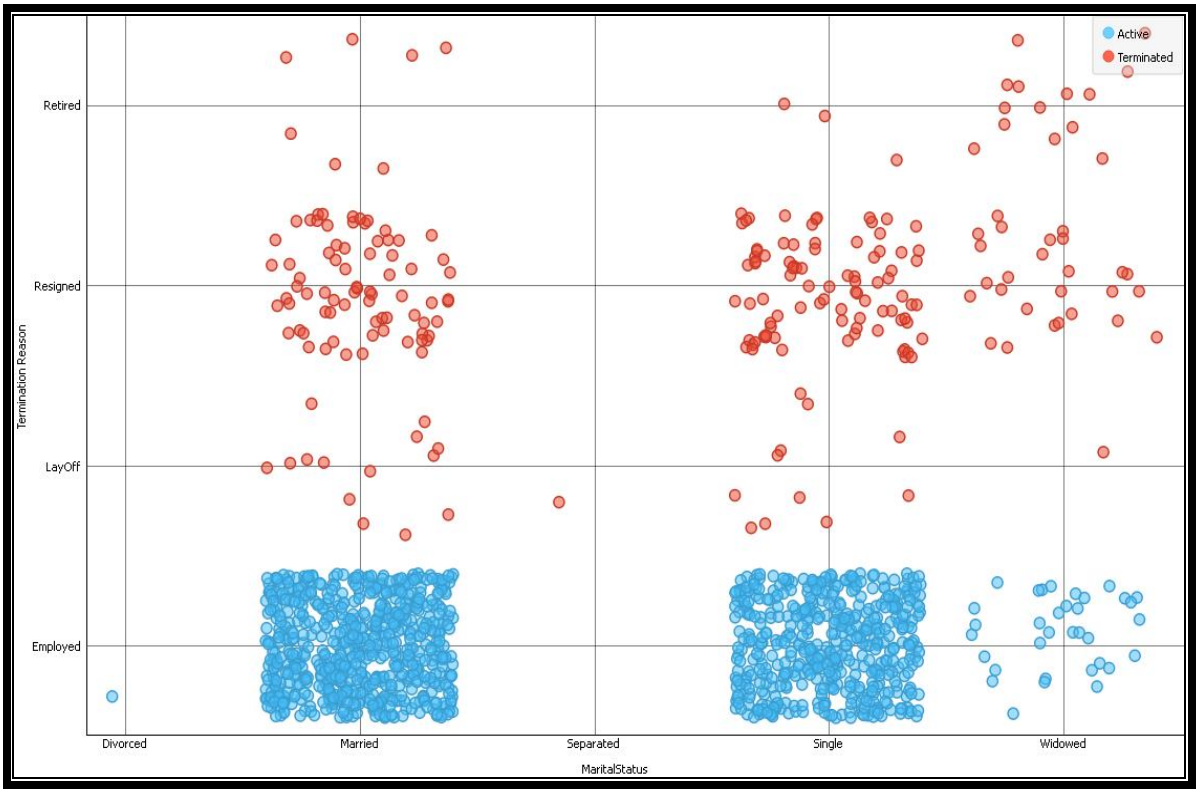


Figure 11. Marital Status

Figure 11 shows the marital status with the most attrition is single, based on the visualization data analysis.

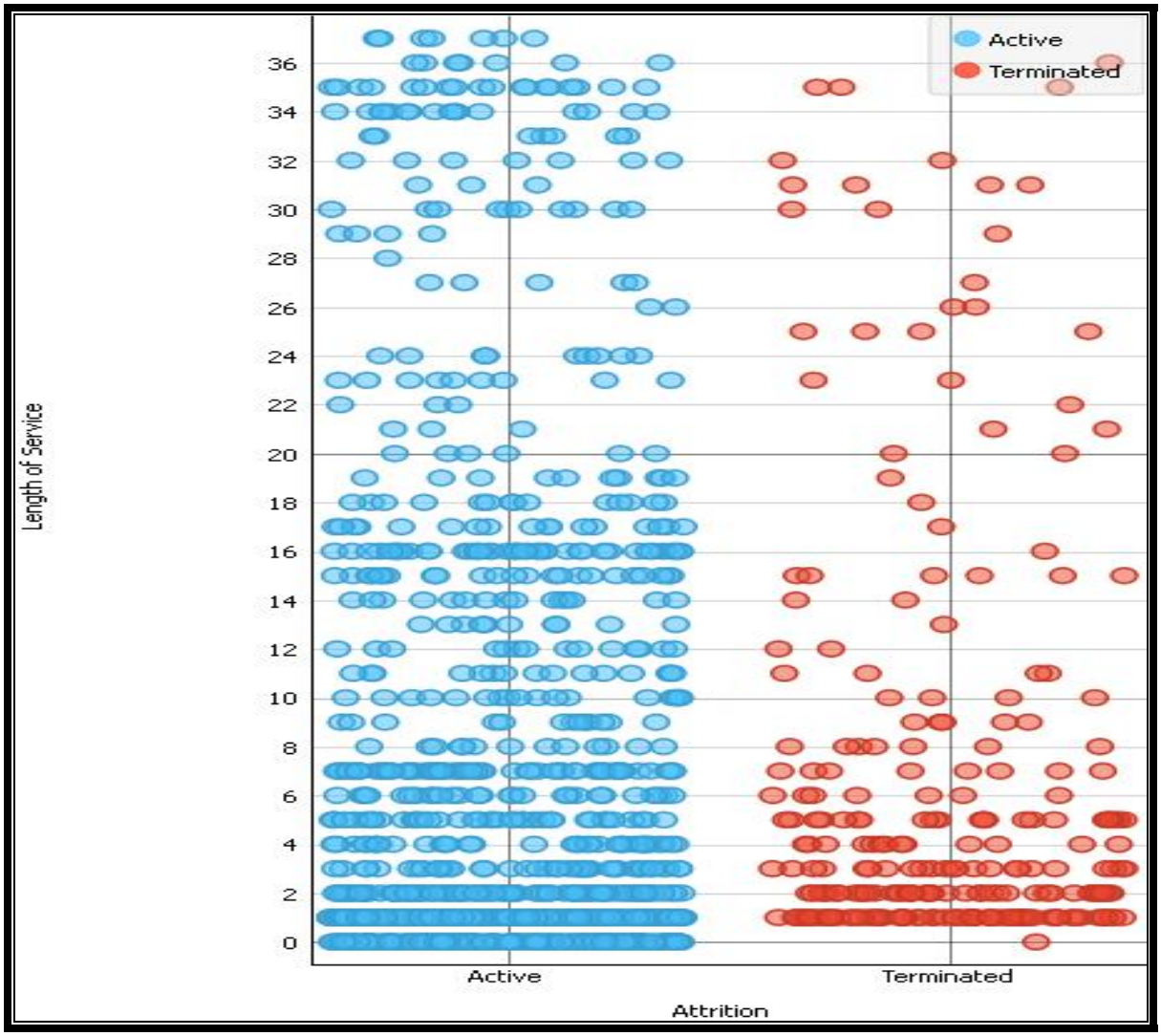


Figure 12. Length of Service

Figure 12 shows the length of service attrition of the selected variables. The graph shows the highest attrition visualized in scatter plot. The distribution reviews the length of service attrition range from 0 to 37 years of service who leave or layoff.

CONCLUSION

Based on the analysed data the researcher derived the following conclusion:

1. As to profile of the employees of Panagsinan State University in terms of age the researchers consider that age most employees leave their job is from the range of 25 to 45 years of age the researcher considered that the age most employer leave their job.
2. In terms of employment status the majority attrition is contractual status and most of them are male, this is expected gender the employee status are contractual because of their status most of them are leave their job this is expected since the employee status be employed at the age of 21 is mandated to retired at 60 to 65 years of age in their service.
3. Employees in their marital status the highest attrition is single, based on the visualization data analysis based on the study the young age are most of them leaving and seek a company for their security tenure.
4. Lastly the length of service some employees leave their job in a few month time while some reaches 37 years of service which is most of them are male and married man.
5. Further, in this paper, a data mining analysis was used to visualize the data from Pangasinan State University. The result on this study will serves as a future reference for other researchers who will conduct the similar or comparable study.

REFERENCES

[1] Nadine Kammerlander (n.d), Organizational Adaptation to Discontinuous Technological Change: The Effects of Family Influence and Organizational Identify. Retrieved from: http://c.ymcdn.com/sites/www.ffi.org/resource/resmgr/FFI_on_Friday/Organizational_Adaptation.pdf

[2] http://journals.upd.edu.ph/index.php/pmr/article/viewFile/2791/2605

[3] Supangco, Vivien T. (2012), Strategic HR Practices in Some Organizations in the Philippines, Philippine Management Review 2012, Vol. 19, 35-48. Retrieved from:

[4] Elfers, A.M., Plecki, M. and Knapp, M. S. (2006). Teacher mobility: Looking more closely at othe movers within a state system. Peabody Journal of Education, 81(3), 94-127.

[5] Henke, R., Zahn, L. and Carroll, C. (2001). Attrition of new teachers among recent college graduates: Comparing occupational stability among 1992–1993 college graduates who taught and those who worked in other occupations. Education Statistics Quarterly, 3(2), 69-76.

[6] McKee, K. L. (2003). A Study Of Factors Related To Teacher Retention. Thesis Prepared For The Degree Of Master of Science, University Of North Texas.

[7] Barmby, P. (2006). Improving teacher recruitment and retention: the importance of workload and pupil behavior. Educational Research, 48(3), 247-265.

[8] Guarino, C. M., Santibanez, L. and Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. Review of Educational Research, 76(2), 173-208.

[9] Azuka, B. F. (2003). The challenges of mathematics in Nigeria economic and technical development – implication for tertiary education. Abacus: J. Math. Assoc. Niger. 28(1): 18-26

[10] Ingersoll, R. (2003). Is there really a teacher shortage? Seattle, WA: Center for the Study of Teaching and Policy

[11] Anderson, L. and Olsen, B. (2005). Studying the career pathways of urban teachers in Los Angeles: Who stays, who leaves, who shifts into other urban education work? Los Angeles: Institute for Democracy, Education and Access, Center X UCLA

[12] Stockard, J. and Lehman, M. B. (2004). Influences on the satisfaction and retention of 1st-year teachers: The importance of effective school management. Education Administration Quarterly, 40,

[13] Boyd, D., Lankford, H., Loeb, S. and Wyckoff, J. (2006). How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement. Education Finance and Policy, 1(2), 176-216.

[14] Bobbitt, S., Leich, M., Whitener, S. and Lynch, H. (1994). Characteristics of stayers, movers, and leavers: Results from the teacher follow up survey, 1991-92. Washington, DC: National Center for Education Statistics.

[15] Dinham, S. (1994). Enhancing the quality of teacher satisfaction. Paper presented at the national Conference of the Australian College of Education, Launceston, Tasmania, Australia.

[16] Boyd, D., Lankford, H., Loeb, S. and Wyckoff, J. (2006). How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement. Education Finance and Policy, 1(2), 176-216.

[17] Long, C. S., M. A. Ajagbe, Khalil Md Nor and E. S. Suleiman (2012b). The Approaches to Increase Employees' Loyalty: A Review on Employees' Turnover Models. Australian Journal of Basic and Applied Sciences, 6(10): 282-291, 2012.

[18] Azuka, B. F. (2003). The challenges of mathematics in Nigeria economic and technical development – implication for tertiary education. Abacus: J. Math. Assoc. Niger. 28(1): 18-26

[19] Nicholas Walliman (2011), Research Method: The Basic, Retrieved from: https://edisciplinas.usp.br/pluginfile.php/2317618/mod_resource/content/1/BLOCO%202_Research%20Methods%20The%20Basics.pdf

★★★