Data Analytics: An Exploration of Quality Control to Determine Students' Academic Performance

Oluwafemi Samson BALOGUN
School of Computing, University of Eastern Finland, FI-70211, Finland
samson.balogun@uef.fi*

Sunday Adewale OLALEYE
Department of Marketing, Management and International Business, Oulu Business School, Erkki Koiso-Kanttilan Katu, Oulu 90570, Finland
sunday.olaleye@oulu.fi

Richard Osei AGJEI
School of public health, Imperial College, London
richardagjei65@gmail.com

Akwasi Gyamerah ADUSEI
Department of Industrial Engineering and Management, University of Oulu, Product Management, Erkki Koiso-Kanttilan Katu, Oulu 90570, Finland
akwasi.adusei@oulu.fi

Abstract

Quality control and improvement is a crucial process development of any institution that craves growth. One part of the SPC approach is to aid the constant improvement of performance by further reducing unexplained variability. Another aspect of Statistical Process Control (SPC) is that planned and unplanned changes signaled as fast as considering the natural process variability. This paper aimed to determine whether students' performance is significantly distributed according to academic patterns using the quality control procedure. This study found that one of the notable Nigerian Private University student academic performances drawn from three engineering departments based on the mean chart is in control and out of control, indicating excellent, intermediate, and lower results. The study also shows upper, average, and lower results with a close margin. This insight is an interdepartmental issue. The school managers need to formulate a holistic policy that will improve the existing academic performance to move the outlier students from worst to better and from better to best.

Keywords: Academic performance, Statistical Process Control, Control Chart, Assignable cause, GPA

Introduction

Quality control and improvement is a crucial process development of any institution that craves growth. Aslam, Saghir, and Ahmad (2020) categorize quality improvement as an intention for rigorous competition and a desire for profit maximization. People believe that statistical process control (SPC) is peculiar to the manufacturing processes. It is still a growing process in health, education, government services, Internet traffic monitoring, and environment protection (Qiu, 2019). There is a caution by earlier study not to tamper with the existing process until it is statistically proving that the process is misbehaving (Aslam et al. 2020). More initial research by Elepo and Balogun (2016) has established a misbehaving process in Nigerian University student academic performance, hence calls for this follow-up study.

Statistical Process Control (SPC) is a structure of thoughts (Wheeler and Chambers, 1992) and a collection of problem-solving tools useful in accomplishing process stability through reduced variability (Quesenberry, 1997; Mitra, 1998; Montgomery, 2009). One part of the SPC approach is to aid the constant improvement of performance by further reducing unexplained variability. Another aspect of SPC is planned and unplanned changes that signal as fast as considering the natural process variability (De Vries and Reneau, 2010). Statistical Process control began in

Cite this Article as: Oluwafemi Samson BALOGUN, Sunday Adewale OLALEYE, Richard Osei AGJEI and Akwasi Gyamerah ADUSEI "Data Analytics: An Exploration of Quality Control to Determine Students' Academic Performance" Proceedings of the 36th International Business Information Management Association (IBIMA), ISBN: 978-0-9998551-5-7, 4-5 November 2020, Granada, Spain.

the manufacturing industries during the 1920s, yet there are presently bounteous applications in health care and disease transmission (Woodall, 2006; Thor et al. 2007).

According to Carey and Stake (2003), "Statistical Process Control (SPC) is a theory, technique, and collection of approaches for the continuous improvement of systems, procedures, and outcomes. The SPC methodology is based on data learning and focused on the principle of variance (understanding common and special causes). The SPC strategy combines analytical; process thought, avoidance, stratification, stability, capacity, and prediction. SPC involves estimation, data collection techniques, and expected research. Graphical techniques, such as Shewhart charts (more generally referred to as 'control charts'), run maps, frequency graphs, histograms, Pareto analysis, scatter diagrams, and flow diagrams, are the key instruments used in SPCs." The control chart is one of the critical instruments of the SPC. A control chart is a statistical tool used for observations plotted after some time to see if the mechanism is working, as it should (De Vries and Reneau, 2010).

Morales (2013) also defined control charts as apparatuses of Statistical Process Control (SPC) that monitor a production process's condition, identifying when a product change's quality characteristics. The concept of 'control' relates to the quality feature within limits set (control limits) to ensure output and quality efficiency. If the attribute (i.e., weight, length, and dimensions) is not within these limits, then the process is in an "out-of-control" state. In such a case, it is essential to discover and address the assignable cause that originated this state (failure).

This paper aims to determine whether or not students' performance is significantly distributed according to academic patterns using the quality control procedure, detecting any statistically significant positive or negative shift in a student's GPA from the desired target level using quality control charts.

This study will facilitate proper monitoring of student performance in the university so that "non-conforming" scores are identified quickly so that such students could properly and adequately be advised to forestall poor outcomes shortly.

Recent studies examine the strength of academic performance agreement as a counseling guide for the University (Balogun, Moshin and Olaleye, 2020). Other studies also applied SPC to monitor mechanized sugarcane harvesting (Paixão et al., 2020) and the monitoring of anatomical changes of individual patients during the head-and-neck radiotherapy process (Lowther et al. 2019). Integrated production maintains a policy for unreliable manufacturing systems (Bahria, Chelbi, Bouchriha, and Dridi, 2019). The proposal of an intelligent SPC method based on feature learning (Zan et al. 2020) and Qiu (2019) highlights some recent literature that showcases nonparametric SPC, control charts for monitoring Spatio-temporal and dynamic processes. These studies portray the vast research domain of SPC. This study divides into six sections. The first section captures the introduction, the second section with literature, the third section with methodology, the fourth section with results, the fifth section with discussion, and the last section concludes the study.

Literature Review

Predicting Student's Academic Performance

Predicting students' academic performance is a crucial part of institutions of higher learning. A better appreciation of the factors that influence students' academic performance is a challenging research exercise because of many diverse aspects such as previous academic performance, interaction with teachers, cultural, and social (Romero and Ventura, 2013). Numerous studies have been carried out on these factors, and they had brought about concrete results. Some works have been done on how socioeconomic status affects students' academic performance (Goddard, Sweetland, and Hoy, 2000). Some other researchers studied the relationship between student's academic performance and the behaviors of their parents (Attaway and Bry, 2004) while others focused on the efficiency of the teacher to improve student's academic performance (Bassam, Ahmed, Mohammad, Mohammad, 2014; Gerber and Fin, 2001). Romero and Ventura (2013) suggest that studies conducted on EDM have been applied to web-based education due to the Learning Management System (LMS) (e.g., WebCT Blackboard and Moodle). Providing information about students' academic assessments and the number of times students access teaching resources. This step is very relevant information to predict students' academic performance and discover course weaknesses (Romero, López, Luna, and Ventura, 2013). Prediction and analysis of students' academic performance are essential for student academic progress (Baker, Corbett, and Koedinger, 2004; Romero and Ventura, 2013).

The objective of predicting student academic performance is to estimate the unknown value that describes the student. In education, the values usually predicted are performance, knowledge, score, or mark. The value of a student's academic performance is either continuous/numerical (regression- Draper and Smith, 1998) value or discrete/categorical (classification-Espejo, Ventura, Herrera, 2010) value. Prediction of a student's academic performance remains an old and widespread application of data mining in education. Different techniques and models, such as neural networks, correlation analysis, rule-based systems, regression, and Bayesian networks, have been applied.

Intelligent Tutoring Systems

Intelligent Tutoring Systems (Hämäläinen and Vinni, 2006) is a machine learning method employed to student's academic performance to predict either fail or pass in a course. A similar algorithm, such as Moodle usage data, is utilized to predict students' final marks (Romero, Ventura, Hervás, Gonzales, 2008). The Moodle usage data algorithm is used to predict students' academic final grade dependent on features extracted from logged data (Minaei-Bidgoli, Kashy, Kortmeyer, Punch, 2003) and to predict University students' academic performance (Ibrahim and Rusli, 2007).

Neural Network

Different types of neural network models have been used to predict the performance of a candidate being considered for admission into the university (Oladokun, Adebanjo, Charles-owaba, 2008); predict final student grades (Gedeon and Turner, 1993); performance from test scores

(Fausett and Elwasif, 1994); predict the number of errors a student will make (Want and Mitrovic, 2002); and predict students' marks from Moodle logs (Delgado, Gibaja, Pegalajar, Pérez, 2006). The application of Bayesian networks is seen in the prediction of student's academic performance within a tutoring system (Pardos, Heffernan, Anderson, Heffernan, 2007); student-applicant performance (Haddawy, Thi, Hien, 2007); to predict a future graduate's Cumulative Grade Point Average (CGPA) based on the background of the applicant at the time of admission (Hien and Haddawy, 2007).

Rule-Based Systems

Different types of rule-based systems have been used in the prediction of student's academic performance in terms of predicting the marks in an e-learning environment through the application of fuzzy association rules (Nebot, Castro, Vellido, Mugica, 2006); to predict, monitor, and evaluate student's academic performance by applying rule induction (Ogor, 2007). Rule-based systems can predict students' academic learner performance based on the learning portfolios compiled using key formative assessment rules (Chen, c., Chen, Li, 2007). Rule-based systems can predict student grades in LMSs, using grammar guided genetic programming (Zafra and Ventura, 2009). Using decision tree rule-based systems can be applied in predicting students' academic performance and providing well-timed lessons in web-based eLearning systems (Chan, 2007).

Regression Techniques

Several regression techniques, such as locally weighted linear regression, neural networks, support vector machines, model trees, and linear regression, have been applied to predict students' marks in an open university (Kotsiantis and Pintelas, 2005). Linear regression prediction models can predict end-of-year accountability assessment scores (Anozie and Junker, 2006). In predicting students' performance from the log and test scores in web-based instruction, a multivariable regression model technique has been employed (Yu, Jannasch-pennell, Digangi, Wasson, 1999). Similarly, stepwise linear regression can predict students' academic performance (Golding, Donalson, 2006). Using multiple regression factors that have the probability of predicting student's academic success in courses run in colleges can be identified (Martinez, 2001), while linear regression can be used to predict student's examination results in courses taught in distance education (Myller, Suhonen, Sutinen, 2002). Again, using a robust Ridge regression algorithm, the linear regression technique is often applied in predicting the probability a student has of producing the right answer to a problem in an ITS (Cetintas, Xin, Hord, 2009). To predict end-of-year accountability assessment scores (Anozie & Junker, 2006) and indicate a student's academic performance test score (Feng, Heffernan, Koedinger, 2005).

Bayesian Networks

Bayesian networks predict student knowledge through cognitive tutors (Baker, Corbett, Aleven, 2008). Bayesian networks have been used to detect students' academic learning styles in a web-based education system (Garcia, Amandi, Schiaffino, Campo, 2007). To correctly predict whether a student will answer a problem correctly, Bayesian networks can be applied (Jonsson, Hasmik, Johns, Mehranian, Arroyo, Woolf, Barto, Fisher, Mahadevan, 2005). Bayesian networks can be used to model a student's changing state of knowledge within the period of skill acquisition in ITS (Chang, Beck, Mostow, Corbett, 2006). Bayesian networks are used to infer intangible learning variables from students' academic performance help-seeking behavior, particularly in a web-based tutoring system (Arroyo, Murray, Woolf, 2004) and for knowledge tracing to authenticate the impact of self-discipline on students' academic knowledge and learning (Gong, Rai, Beck, Heffernan, 2009).

Methodology

Data for this study were collected from students' academic records in the Department of Information and Communication Engineering, Mechanical Engineering and Petroleum Engineering, Covenant University, Ota, Ogun State, Nigeria from 2004/2005, 2005/2006, 2006/2007, 2007/2008, 2008/2009, and 2009/2010 academic sessions. Thus, this study's data is secondary data of 195, 166, and 195 observations on students' G.P.A. for six academic sessions. In this work, we will be using a control chart to monitor the mean and the variability of the process (students' academic performance). The analysis is carried out on IBM SPSS version 25.

Derivation of $\bar{x}$ and s formula

According to Montgomery (2009) setting up and operating $\bar{x}$ *and* s control charts require a similar sequence of steps as that of $\bar{x}$ *and* R chart, except that for each of the sample we calculate the sample mean $\bar{x}$ and sample standard deviation s. If the σ^2 defined as the unknown variance of the probability distribution, then the unbiased estimator of σ^2 is as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

However, the sample standard deviations is not an unbiased estimator of σ. If the underlying distribution is normal, s estimates $c_4\sigma$, where c_4 is a constant which depends on the sample size n. Furthermore, the standard deviation of s is $\sigma\sqrt{1-c_4}$. The information so far is used to establish the control chart of $\bar{x}$ *and* s. Considering the case where the standard value is given for σ. Since $E(s) = c_4\sigma$, the center line for the control chart is $c_4\sigma$. Using the three sigma control limits for s are then

$$UCL = c_4\sigma + 3\sigma\sqrt{1-c_4^2}$$

$$LCL = c_4\sigma - 3\sigma\sqrt{1-c_4^2}$$

It is therefore customary to define the two constants as:

$$B_5 = c_4 - 3\sqrt{1-c_4^2} \quad and$$

$$B_6 = c_4 + 3\sqrt{1-c_4^2}$$

Consequently, the parameters of the s chart with a standard value for σ is given as:

$$UCL = B_6\sigma, \qquad CL = c_4\sigma, \qquad LCL = B_5\sigma$$

The values of B_5 and B_6 are from the table with various sample sizes. The statistic $\frac{\bar{s}}{c_4}$ is an unbiased estimator of σ. Therefore, the parameters of the S chart would be

$$UCL = \bar{s} + 3\frac{\bar{s}}{c_4}\sqrt{1-c_4}$$

$$Center\ line(CL) = \bar{s}$$

$$LCL = \bar{s} - 3\frac{\bar{s}}{c_4}\sqrt{1-c_4}$$

The constants are defined as

$$B_3 = 1 - \frac{3}{c_4}\sqrt{1-c_4^2} \quad and$$

$$B_4 = 1 + \frac{3}{c_4}\sqrt{1-c_4^2}$$

Consequently, the parameters of the s chart are written as:

$$UCL = B_4\bar{s}, \quad Center\ line\ (CL) = \bar{s}, \quad LCL = B_3\bar{s}$$

Note that $B_4 = \frac{B_6}{c_4}$ and $B_3 = \frac{B_5}{c_4}$. When $\frac{\bar{s}}{c_4}$ is used to estimate σ, the control limits on the corresponding $\bar{x}$ chart is defined as:

$$UCL = \bar{\bar{x}} + \frac{3\bar{s}}{c_4\sqrt{n}}$$

$$Center\ line\ (CL) = \bar{\bar{x}}$$

$$LCL = \bar{\bar{x}} - \frac{3\bar{s}}{c_4\sqrt{n}}$$

Assume the constants $A_3 = \frac{3}{c_4\sqrt{n}}$. Then the $\bar{x}$ chart parameter becomes:

$$UCL = \bar{\bar{x}} + A_3\bar{s}, \quad Center\ line\ (CL) = \bar{\bar{x}}, \ LCL = \bar{\bar{x}} - A_3\bar{s}$$

Definition of symbols used

According to Akinrefon and Balogun (2015) the symbols used for $\bar{x}\ and\ s$ are defined as follows:

$\bar{\bar{x}}$ = Average of the subgroup average
$\bar{x}$ = Average of subgroup
n = Number of subgroups
σ = Population standard deviation of the subgroup average
s = Sample standard deviation of the subgroup average
$\bar{s}$ = Average of the standard deviation of the subgroup
UCL = Upper Control Limits
LCL = Lower Control Limits
B_4 = Approximation factor used to calculate within subgroup standard deviation chart control limit
B_3 = Approximation factor used to calculate within subgroup standard deviation chart control limit
A_3 = Approximation factor used to calculate control limit
c_4 = Approximation factor used to calculate within subgroup standard deviation

Results

This section interprets the results from mean and standard deviation control charts from three departments, namely: Petroleum Engineering, Chemical Engineering, Information, and Communication Engineering department.
Petroleum Engineering Department

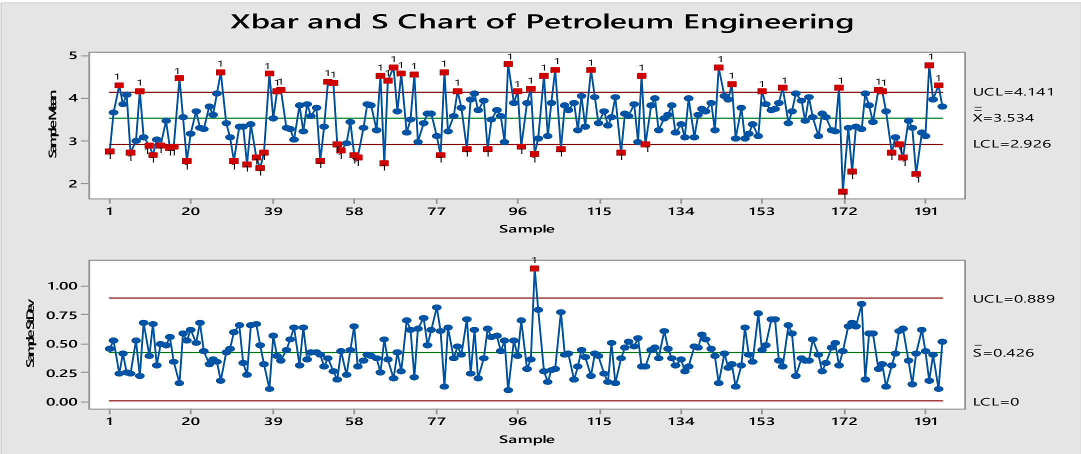

Fig. 1: Xbar and S chart for Petroleum Engineering department

Interpretation of the control chart

For mean chart: $UCL = \bar{\bar{x}} + A_3\bar{s} = 4.141, CL = \bar{\bar{x}} = 3.534\ and\ \ LCL = \bar{\bar{x}} - A_3\bar{s} = 2.926$. The mean chart of the student's GPA from the Petroleum Engineering department shows that not all the points are within the control limits. Thus, the student's performance is out of control, meaning that there is some assignable cause of variation in the student's performance, which should be observed and corrected. This study emphasizes points that fall below the control limit, i.e., students with poor results. From Fig. 1, the upper limit is 4.14, which means that students whose GPA falls on and above the upper control limits are extremely good since those students would graduate with second-class upper and first-class from the department. The department's average student performance is 3.53, which corresponds to the second grade, which means that the average student in the department will graduate with the second-class grade. The lower limit is 2.93, which implies acceptable performance and is thus considered the lower bound for students' academic performance from the department. A GPA, which drops below the limit, is considered out of control. The Students whose GPA falls below the lower limits will be graduating with a second-class, a third class, or a pass. Generally, the results of students graduating from the department are strong, except for a few students whose performance is poor as they graduated in the third grade. This result is due to the facilities available to students and lecturers, the academic program's quality, and the lecturers' attitude. This success is evident in private universities in Nigeria. From standard deviation chart: $UCL = B_4\bar{s} = 0.889, CL = \bar{s} = 0.426\ and\ \ LCL = B_3\bar{s} = 0$. The standard deviation chart of students' GPA from the Petroleum Engineering Department indicates that two-point is outside the control limits. As such, the process is out of control due to an assignable cause of variation in student results. Students whose GPA falls below the lower control limit will graduate with an average or unsatisfactory result. In contrast, those who fall above the upper control limit will graduate with an outstanding result.

Chemical Engineering

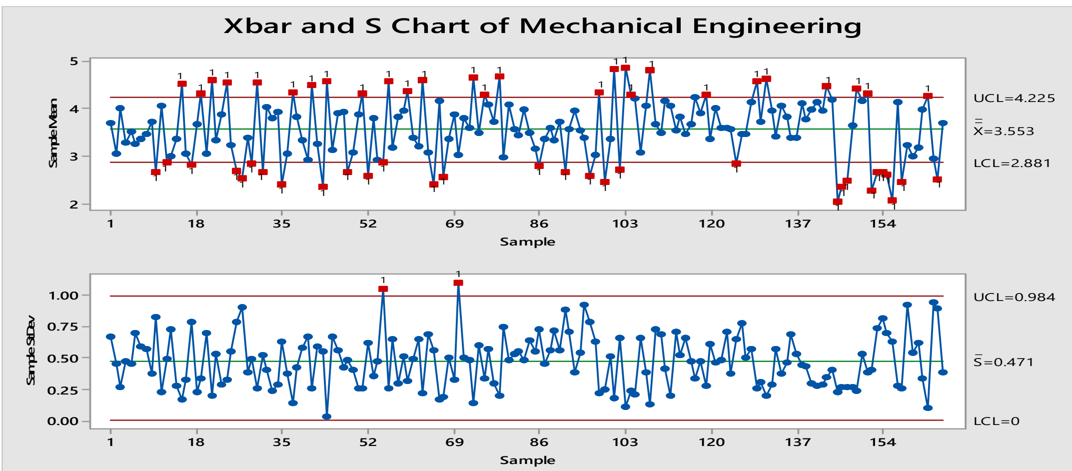

Fig. 2: Xbar and S chart for Chemical Engineering department

Interpretation of the control chart

For mean chart: $UCL = \bar{\bar{x}} + A_3\bar{s} = 4.225, CL = \bar{\bar{x}} = 3.553\ and\ \ LCL = \bar{\bar{x}} - A_3\bar{s} = 2.881$. The mean chart of students' GPA from the Chemical Engineering Department indicates that not all points fall within the control limits. Thus, the student's performance is out of control, meaning that there is some assignable cause of variation in the student's performance, which should be detected and corrected. This study emphasizes points that fall below the control limit, i.e., students with poor results. Fig. 2 shows the upper limits are 4.23, indicating students whose GPA falls above the upper control limits are considered extremely good. Such students would be graduating with second-class upper and first-class grades from the department. The department's average student performance is 3.55, which corresponds to the second grade, which means that the average student in the department will graduate with the second grade. The lower limit is 2.88, which implies acceptable performance and is thus considered the lower bound for students' academic performance from the department. The GPA, which falls below the limit, is considered out of control. Students' GPA falls on the lower limits would be awarded a second class, a third class, or a pass. Generally, the results of students graduating from the department are good, except for a few students whose performance is bad as they graduated from the Third-Class grade.

From standard deviation chart: $UCL = B_4\bar{s} = 0.984, CL = \bar{s} = 0.471\ and\ \ LCL = B_3\bar{s} = 0$. The standard deviation chart of students' GPA from the Department of Mechanical Engineering indicates that one point is outside the control limits. As such, the process is out of control due to an assignable cause of variation in student results. The point that falls below the lower control limit is that students will graduate with an average or poor result. In contrast, the students who fall above the upper control limit will graduate with an outstanding result.
Information and Communication Engineering

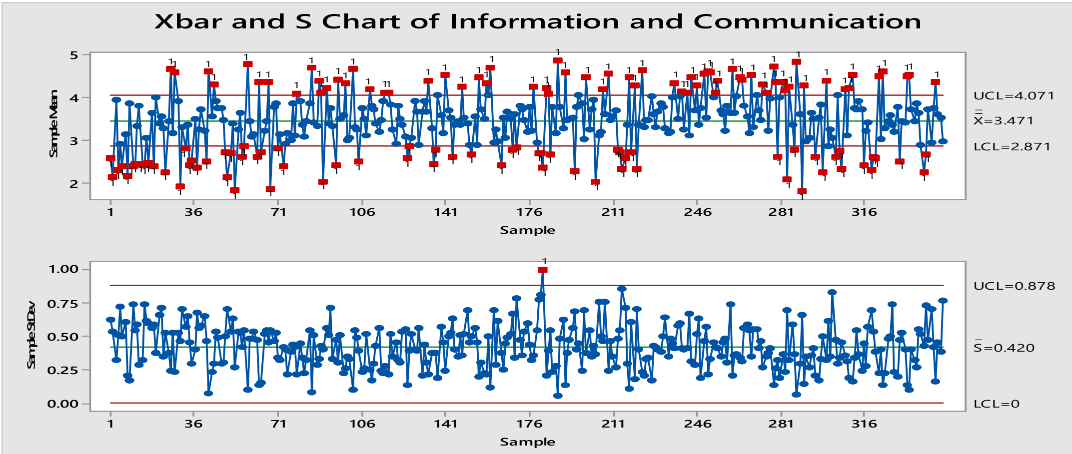

Fig. 3: Xbar and S chart for information and Communication Engineering department

Interpretation of the control chart

For mean chart: $UCL = \bar{\bar{x}} + A_3\bar{s} = 4.071, CL = \bar{\bar{x}} = 3.471\ and\ \ LCL = \bar{\bar{x}} - A_3\bar{s} = 2.871$. The mean chart of the students' GPA from the Information and Communication Engineering Department indicates that not all points are within the control limits. Thus, the student's performance is out of control, meaning that there is some assignable cause of variation in the student's performance, which should be detected and corrected. This study emphasizes points that fall below the control limit, i.e., students with poor results. Fig. 3 shows the upper limit is 4.07, implying that students whose GPA falls below and above the upper control limits are extremely good since those students would graduate with a second-class upper and first-class grade from the department. The average students' performance in the department is 3.47, which corresponds to the lower grade second class, which means that the average student in the department will graduate with a second-class lower grade. The lower limit is 2.87, which implies acceptable performance and is thus considered the lower bound for students' academic performance. A GPA, which falls below the limit, is considered out of control. Students' GPA falls within the lower limits will be graduating with a second-class lower grade, a third class, or a pass. The results of students graduating from the department are good, except for a few students whose performance is poor as they graduated from the Third-Class grade. From standard deviation chart: $UCL = B_4\bar{s} = 0.878, CL = \bar{s} = 0.420\ and\ \ LCL = B_3\bar{s} = 0$. The standard deviation chart of students' GPA from the Information and Communication Engineering Department indicates that one point is outside the control limits. As such, the process is out of control due to an assignable cause of variation in student results. The point that falls below the lower control limit is that students will graduate with an average or poor result. In contrast, those who fall above the upper control limit will graduate with an outstanding result.

Discussion

This study found that one of the notable Nigerian Private University student academic performances drawn from three engineering departments based on the mean chart is in control and out of control, indicating excellent, intermediate, and lower results. The study also shows upper, average, and lower results with a close margin. The clustering of ideal, middle, and lower points is distinct but only requires the student to get 0.62 to cross from lower to average and 0.59 from average to the upper level in the GPA SPC chart. Comparatively, the Petroleum Engineering department emerged in the upper point, became second in average point, and first in a lower point. Chemical Engineering attained second in upper point, first in average point, and second in lower point while Information and Communication Engineering came third in upper, average, and lower points. Petroleum Engineering and Chemical Engineering department undulate while Information and Communication Engineering maintain a steady rank in all the applied metric system. This study's result is similar to Haddawy, Thi, and Hien (2007), which predicted the future graduate's Cumulative Grade Point Average (CGPA) with the Neural Network data analysis technique. However, the difference between the upper, average, and lower points of each engineering department in a private setting is not conspicuous. Additional research is needed to determine how students who fall into the lower category can improve and become average students and, likewise, for intermediate students to become upper students. Excellent results in the private university will enhance public relations and boost students' chances of getting hired after their studies. Producing exceptional students in a Private University may also be an extension of reputation management, competitive advantage, and indirectly for-profit maximization (Olaleye, Sanusi, and Salo, 2018; Aslam, Saghir, and Ahmad, 2020).

Conclusion

This study found a considerable variation in the students' academic performance, especially the out of control groups. This insight is an interdepartmental issue. The school managers need to formulate a holistic policy that will improve the existing academic performance to move the outlier students from worst to better and from better to best. This policy will pave the way for continuous student academic improvement. This study is significant because it is foremost to apply SPC to determine private university students' academic performance. Despite this study's impact, it is limited to only three departments, which may not represent the entire private university's programs. It will be interesting if future studies can study the whole departments established in a private university concerning students' academic performance. Future studies also can compare private and public universities' student academic performance in emerging countries.

References

- Anozie, N., Junker, B.W. (2006). Predicting end-of-year accountability assessment scores from monthly student records in an online tutoring system. In Educational Data Mining AAAI Workshop, California, 1-6.
- Arroyo, I., Murray, T., Woolf, B.P. (2004). Inferring Unobservable Learning Variables from Students' Help Seeking Behavior. In International Conference on Intelligent Tutoring System, Brazil, 782-784.
- Aslam, M., Saghir, A., & Ahmad, L. (2020). Introduction to Statistical Process Control. John Wiley & Sons.
- Attaway, N. M., & Bry, B. H. "Parenting style and black adolescents' academic achievement," Journal of Black Psychology, 30, pp.229–247, 2004.
- Bahria, N., Chelbi, A., Bouchriha, H., & Dridi, I. H. (2019). Integrated production, statistical process control, and maintenance policy for unreliable manufacturing systems. International Journal of Production Research, 57(8), 2548-2570.
- Baker, R., Corbett, A.T., Aleven, V. (2008). Improving contextual models of guessing and slipping with a truncated training set. In International Conference on Educational Data Mining, Montreal, Canada, 67-76.
- Baker, R.S., Corbett, A.T. & Koedinger, K.R. "Detecting Student Misuse of Intelligent Tutoring Systems," Proceedings of the 7th International Conference on Intelligent Tutoring Systems, pp.531-540, 2004.
- Balogun, O.S., Moshin, M. & Olaleye, S.A. (2020). The strength of agreement of students' academic performances as a counseling guide for the University prospective Admission seekers. In Proceedings of the 35th International Business Information Management Association Conference (IBIMA) (pp. 11935-11945).
- Bassam Zafar, Ahmed Mueen, Mohammad Awedh, Mohammad Balubaid, "Game-based learning with native language hint and their effects on student academic performance in a Saudi Arabia community college" Computer in Education vol. 1, no. 4, pp. 371-384, 2014.
- Carey, R. G., & Stake, L. V. (2003). Improving healthcare with control charts: basic and advanced SPC methods and case studies. ASQ Quality Press.
- Cetintas, A., Si, L., Xin, Y.P., Hord, C. (2009). Predicting correctness of problem solving from low-level log data in intelligent tutoring systems. In International Conference on Educational Data Mining, Cordoba, Spain, 230-238.
- Chan, C.C. (2007). A Framework for Assessing Usage of Web-Based e-Learning Systems. In International Conference on innovative Computing, Information and Control, Washington, DC, 147- 151.
- Chang, K.M., Beck, J.E., Mostow, J. & Corbett, A. (2006). A Bayes Net Toolkit for Student Modeling in Intelligent Tutoring Systems. In International Conference on Intelligent Tutoring Systems, Jhongli, Taiwan, 104-113.
- Chen, C., Chen, M. & Li, Y. (2007). Mining key formative assessment rules based on learner port files for web-based learning systems. In IEEE International Conference on Advanced Learning Technologies, Japan, 1-5.
- De Vries, A., & Reneau, J. K. (2010). Application of statistical process control charts to monitor changes in animal production systems. Journal of Animal Science, 88(suppl_13), E11-E24.
- Delgado, M., Gibaja, E., Pegalajar, M.C., Pérez, O. (2006). Predicting Students' Marks from. Moodle Logs using Neural Network Models. In International Conference on Current Developments in Technology-Assisted Education, Sevilla, Spain, 586-590.
- Draper, N.R. & Smith, H. (1998). Applied Regression Analysis. Wiley.
- Elepo, T.A. & Balogun, O.S. (2016) 'Statistical Analysis on Students' Performance', Covenant Journal of Informatics and Communication Technology, 4(1) Available at: http://journals.covenantuniversity.edu.ng/index.php/cjict/article/view/217 (Accessed: 14 May 2020).

- Espejo, P., Ventura, S. & Herrera, F. (2010). A Survey on the Application of Genetic Programming to Classification. IEEE Transactions on Systems, Man, and Cybernetics-Part C. 40, 2, 121-144.
- Fausett, L.V. & Elwasif, W. (1994). Predicting performance from test scores using backpropagation and counterpropagating. In IEEE World Congress on Computational Intelligence, Paris, France, 3398–3402.
- Feng, M., Heffernan, N. & Koedinger, K. (2005). Looking for sources of error in predicting student's knowledge. In: AAAI'05 workshop on Educational Data Mining, 1-8.
- Garcia, P., Amandi, A., Schiaffino, S. & Campo, M. (2007). Evaluating bayesian networks' precision for detecting student's learning styles. In Computer & Education Journal. 49, 794-808.
- Gedeon, T.D. & Turner, H.S. (1993). Explaining student grades predicted by a neural network. In International conference on Neural Networks, Nagoya, 609-612.
- Gerber, S. B., & Fin, J. D. (2001). "Teacher aides and students" academic achievement Educational Evaluation and Policy Analysis, 23(2), pp.123–143.
- Goddard, R. D., Sweetland, S. R., & Hoy, W. K. (2000). "Academic emphasis of urban elementary schools and student achievement in reading and mathematics: A multilevel analysis," Educational Administration Quarterly, 36(5), pp.683–702.
- Golding, P. & Donalson, O. (2006). Predicting Academic Performance. In Frontiers in Education Conference. San Diego, California, 21-26.
- Gong, Y., Rai, D., Beck, J.E. & Heffernan, N.T. (2009). Does Self-discipline impact students' knowledge and learning? In International Conference on Educational Data Mining, Cordoba, Spain, 61-70.
- Haddawy, P., Thi, N. & Hien, T.N. (2007). A decision support system for evaluating international student applications. In Frontiers In Education Conference, Milwaukee, 1-4.
- Hämäläinen, W. & Vinni, M. (2006). Comparison of machine learning methods for intelligent tutoring systems. In international conference in intelligent tutoring systems, Taiwan, 525-534.
- Hien, N.T.N. & Haddawy, P. (2007). A decision support system for evaluating international student applications. In Frontiers In Education Conference, Milwaukee, 1-6.
- Ibrahim, Z. & Rusli, D. (2007). Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression. In Annual SAS Malaysia Forum, Kuala Lumpur, 1-6.
- Jonsson, A., Hasmik, J. Johns, Mehranian, H., Arroyo, I., Woolf, B., Barto, A., Fisher, D. & Mahadevan, S. (2005). Evaluating the Feasibility of Learning Student Models from Data, In Educational Data Mining AAAI Workshop, Pittsburgh, 1-6.
- Kotsiantis, S.B. & Pintelas, P.E., (2005). Predicting Students' Marks in Hellenic Open University. In IEEE international Conference on Advanced Learning Technologies, Washington, DC, 664-668.
- Lowther, N. J., Hamilton, D. A., Kim, H., Evans, J. M., Marsh, S. H., & Louwe, R. J. (2019). Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy. Physics and Imaging in Radiation Oncology, 9, 21-27.
- Martinez, D., (2001). Predicting Student Outcomes Using Discriminant Function Analysis. In Meeting of the Research and Planning Group, Lake Arrowhead, CA, 1-22.
- Minaei-Bidgoli, B., Kashy, D.A., Kortmeyer, G. & Punch, W.F. (2003). Predicting student performance: an application of data mining methods with an educational Web-based system. In International Conference on Frontiers in Education, 13-18.
- Mitra, A. (1998). Fundamentals of Quality Control and Improvement. 2nd ed. Prentice Hall, Upper Saddle River, NJ.
- Montgomery, D. C. (2009). Introduction to Statistical Quality Control. 6th ed. John Wiley and Sons, New York, NY.
- Morales, S. O. C. (2013). Economic statistical design of integrated X-bar-S control chart with preventive maintenance and general failure distribution. PloS one, 8(3).
- Myller, N., Suhonen, J. & Sutinen, E. (2002). Using Data Mining for Improving Web-Based Course Design. In International Conference on Computers in Education, Washington, 959- 964.
- Nebot, A., Castro, F., Vellido, A. & Mugica, F. (2006). Identification of fuzzy models to predict student's performance in an e-learning environment. In International Conference on Web-based Education, Puerto Vallarta, 74-79.
- Ogor, E.N. (2007). Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques. In Electronics, Robotics and Automotive Mechanics Conference, Washington, DC, 354-359.
- Oladokun, V.O., Adebanjo, A.T. & Charles-owaba, O.E. (2008). Predicting student's academic performance using artificial neural network: A case study of an engineering course. In Pacific Journal of Science and Technology, 9,1, 72-79.
- Olaleye, S. A., Sanusi, I. T., & Salo, J. (2018). Sentiment analysis of social commerce: a harbinger of online reputation management. International Journal of Electronic Business, 14(2), 85-102.

- Paixão, C. S., Voltarelli, M. A., Silva, R. P. D., Borba, M. A. D. P., & Torres, L. S. (2020). Statistical Process Control Applied to Monitor Losses in the Mechanized Sugarcane Harvesting. Engenharia Agrícola, 40(4), 473-480.
- Pardos, Z., Heffernan, N., Anderson, B. & Heffernan, C. (2007). The Effect of Model Granularity on Student Performance Prediction Using Bayesian Networks. In International Conference on User Modeling, Corfu, Greece, 435-439.
- Qiu, P. (2019). Some recent studies in statistical process control. In Statistical quality technologies (pp. 3-19). Springer, Cham.
- Quesenberry, C. P. (1997). SPC Methods for Quality Improvement. John Wiley and Sons, New York, NY.
- Romero, C., López, M. I., Luna, J. M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. Computers & Education, 68, 458-472.
- Romero, C., & Ventura, S. (2013) "Data mining in Education," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(1), pp.12–27.
- Romero, C., Ventura, S., Hervás, C. & Gonzales, P. (2008). Data mining algorithms to classify students. In International Conference on Educational Data Mining, Montreal, Canada, 8-17.
- Thor, J., Lundberg, J., Ask, J., Olsson, J., Carli, C., Härenstam, K. P., & Brommels, M. (2007). Application of statistical process control in healthcare improvement: systematic review. BMJ Quality & Safety, 16(5), 387-399.
- Want, T., Mitrovic, A. (2002). Using Neural Networks to Predict Student's Performance. In International Conference on Computers in Education, Washington, DC, 1-5.
- Wheeler, D. J., & Chambers, D. S. (1992). Understanding Statistical Process Control. 2nd ed. SPC Press, Knoxville, TN.
- Woodall, W. H. (2006). The use of control charts in healthcare and public-health surveillance. J. Qual. Technol. 38:89–104.
- Yu, C.H., Jannasch-pennell, A., Digangi, S. & Wasson, B. (1999). Using On-line interactive statistics for evaluating Web-based instruction, In Journal of Educational Media International, 35, 157-161.
- Zafra, A. & Ventura, S. (2009), Predicting student grades in learning management systems with multiple instance programming. In International Conference on Educational Data Mining, Cordoba, Spain, 307-314.
- Zan, T., Liu, Z., Su, Z., Wang, M., Gao, X., & Chen, D. (2020). Statistical Process Control with Intelligence Based on the Deep Learning Model. *Applied Sciences*, *10*(1), 308.