

Data Analytics in an Open and Distance Learning (ODL) Institution: The case of Botswana Open University (BOU)

A paper to be presented at the Tenth Pan-Commonwealth Forum on Open Learning (PCF10)

Themed: "*Innovations for Educational resilience*":

Organised by the Commonwealth of Learning and Athabasca University

Calgary, Canada, 14 - 16 September 2022.

Moses Nkgari

Botswana Open University (BOU)

Private Bag BO 187, Gaborone,

Botswana

e-mail: mnkgari@staff.bou.ac.bw

Refilwe T. Modikwagae

Botswana Open University (BOU)

Private Bag BO 187, Gaborone,

Botswana

e-mail: rmodikwagae@staff.bou.ac.bw

ABSTRACT

The use of Information and Communication Technologies (ICTs) in ODL allows for the generation of huge amounts of data, sometimes known as "Big data," when applicants and enrolled learners interact with the online platforms. It is argued that; this vast data is captured and stored, but it is often insufficiently analysed to assist learners and the institution in the service delivery or data-driven decision-making processes. This article investigates how this data may be analysed and used to develop strategies to enhance learner services and service delivery in an ODL, specifically at BOU. The report describes the outcomes of a quantitative research study for Botswana Open University, which was done through the analysis of raw data from the BOU Student Management System. This paper offers insights generated from BOU data analytics that might be used to improve the University's teaching and learning. According to BOU's case study, ODL institutions might benefit from data analytics adoption by identifying trends and reorganizing the ODL value chain.

Keywords: Analytics, big-data, learner, ODL, decision-making, data mining

INTRODUCTION

Botswana Open University (BOU) is an Open and Distance Learning (ODL) institution that was a transition from a distance and learning college. The University was established to meet the need for distance learning in higher education in Botswana. BOU offers programs from certificate level to postgraduate level across various schools, otherwise commonly referred to as faculties which are school of education, school of social sciences, school of business and management as well as school of science and technology. The programs were initially offered in a blended mode through which both virtual and face-to-face interactions were utilized, however, since the Covid-19 pandemic, the University has transitioned to virtual delivery mode. This has contributed to the massive data production across BOU systems (learning management systems, student management systems, library systems). However, the researchers of this paper realized that there are more analytics activities that could help draw insights from this abundant data with the intention to inform decision making. Data science has been defined as a “multidisciplinary field that involves a number of methods and techniques to collect data from different sources, then process it in such a way that we can extract some meaningful insights from it to solve real-world problems” (Team, 2021:1). The multidisciplinary element of it allows it to also be used in the education sector to influence decision making for improved service and delivery.

The main objective of this study is to mine and analyse the application’s data in order to observe trends in applications through admission. The insights drawn from this analysis are presented below, followed by recommendations on data application practices that the University could adapt for daily decision making.

RELATED STUDIES

A study was undertaken by Nandeshwar and Chaudhari (2014) to build models they could use in predicting student enrolment. They used admissions data from spring 1999 to fall 2006 which amounted to 28000 total applications. Their main findings were that financial assistance was a contributing factor in students enrolling on their University. This meant that, for increased enrolment, the university could look into increasing its funding of students.

Another similar study was a data mining procedure that was used to select students for enrolment in a particular program (Yadav and Pal, 2012), which was done at the Department of MCA of V.B.S Purvanchal University, Jaunpur, India. The study used a classification method (Decision Tree) to make the predictions. The predictions made were that 33.33% of students enrolled in MCA were likely to obtain Third Class while 50% of them could fail. The observation made was that students with a strong background in mathematics performed better in this program compared to those with just B.As.

(Mulugeta and Borena, nd) undertook a study in the Higher Education Students` Enrolment Forecasting System Using Data Mining Application-Ethiopia. The aim of the study was to develop a predictive model to be used in determining enrolment at the department level across the University. The researchers took three sets of data to make projections for 3 academic years and implored three different models to make projections, the decision tree, neural network model as well as the Bayesian classifier. The study reports that the neural network model which was great in dealing with the non-linear relationship between the predictor and the target variable, was the best model, in this case, to be used for predictions of enrolment in the University.

The objective of this study is that it looks at the trends in applications, so as to build relationships between variables observed in application data to see trends in who applies and consequently see the contributing factors to enrolment. The similarity of this study and the studies mentioned above is that they all rely heavily on student data.

PROJECT PROCEDURE

Data Mining and Pre-Processing

The researchers used applications data from 2019 to 2021 which made a total of 19,976 applications. The data had 65 variables which were from both the demographic and academic information of the applicants. The pre-processing was performed in the python notebook to check the number of rows and columns in the data, as well as establish the variables (columns) in the data. Some variables that were not regarded as critical in this analysis were dropped from the data leading to only 18 variables. The most critical in removing some variables were the anonymization process in which personally identifiable information was removed from the data. Missing data was also located and treated to avoid misrepresentation in the analysis. The researchers further used python software to identify the outliers in the data and treated them accordingly. The data was then analysed on excel for drawing insights and the results are presented below.

Analysis and Project Findings/Discussion

This section provides various insights into the applications data from BOU.

Overview Of Insights From Data Mining Activities

1. Yearly Applications Analysis:

From both charts below, it is evident that the university was affected by the 2020 pandemic. The rate of applications dropped by over two thousand within the years 2019 and 2020. This was most likely due to the global pandemic, COVID-19. In the year, 2021 the rate of application only increased by a thousand from the previous year, not matching up to the year 2019. This is something the university would want to look at to avoid a further drop in the yearly applications.

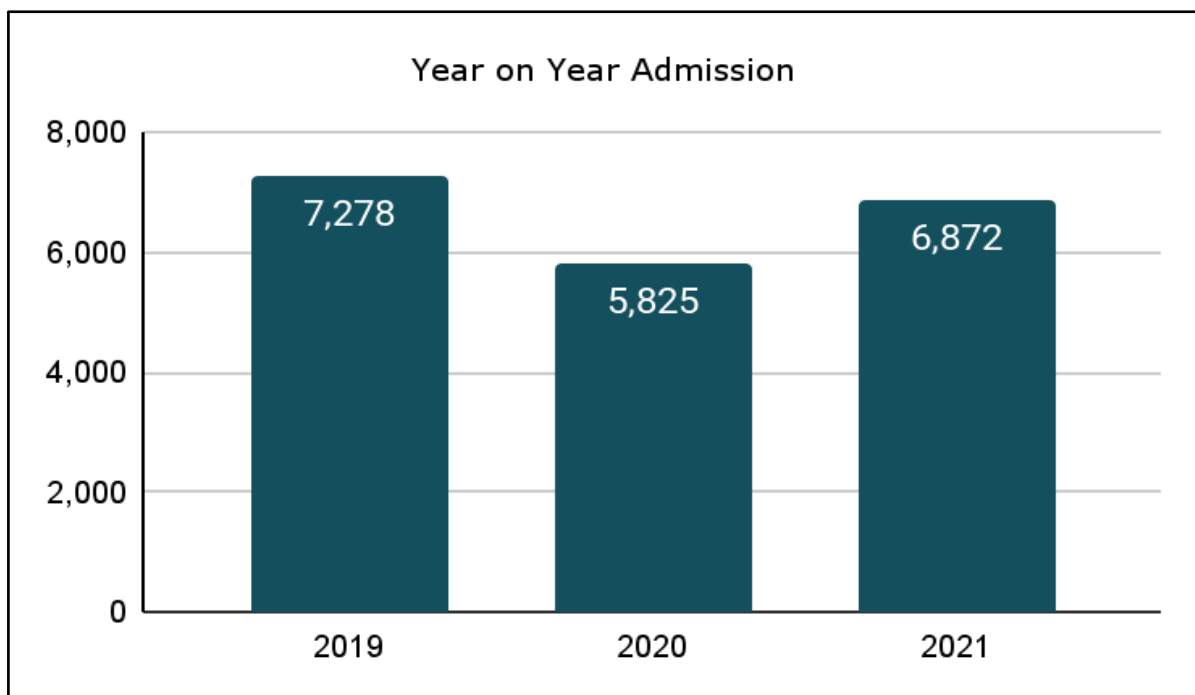


Figure 1: Year-on-Year Admission

Over the years there has been a similar trend for rejection, and acceptance The University seems to maintain a specific range of acceptance rate and rejection rate. The university has maintained an acceptance rate within the range of 63% to 69%, and a rejection rate within the range of 31% to 35%.

From the 7,278 applications received in 2019, 4714 applications were accepted, and 2564 applications were rejected. Of the 5825 applications received in 2020, a total of 3715 were accepted while 2110 of them were rejected. While in 2021, it shows that a total of 4719 applicants were admitted and 2153 rejected from the initial 6872 applications received.

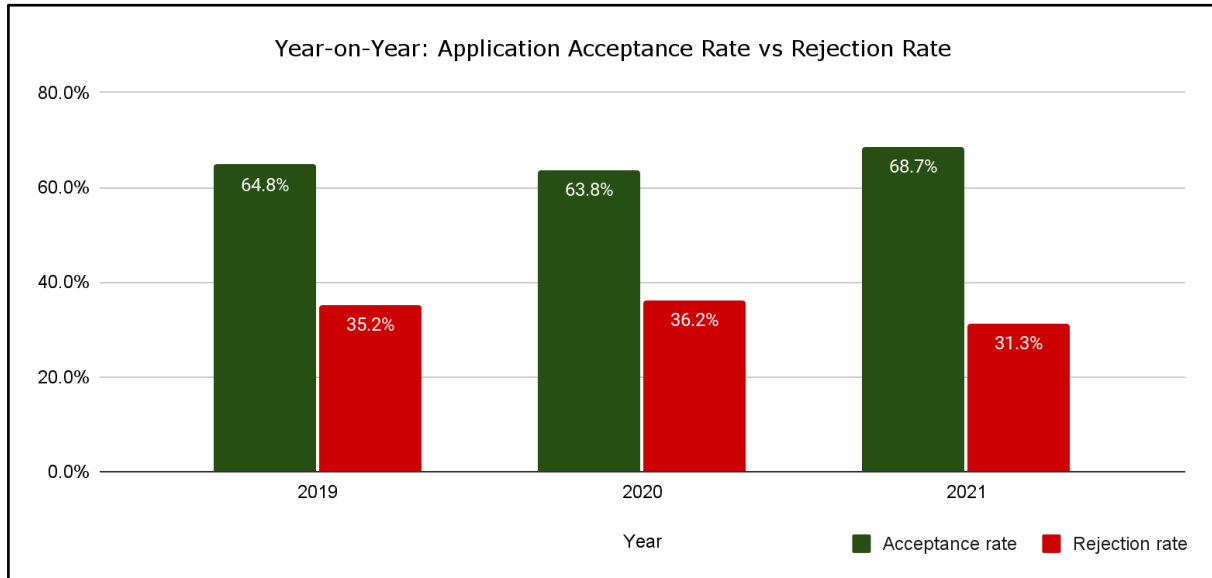


Figure 2: Year-on-Year Application Acceptance vs Rejection Rate

2. Gender Analysis:

The University has had more female than male applicants over the years. This is an interesting trend and we decided to investigate further. The University has done excellently by creating a diverse environment and has encouraged female applicants over the years. The university can leverage this strength to encourage more female applicants. To create a balance between the gender gap, the university should work on attracting more male applicants. With the above found during data mining, further querying of the data shows that there is a higher admission acceptance rate for females compared to males.

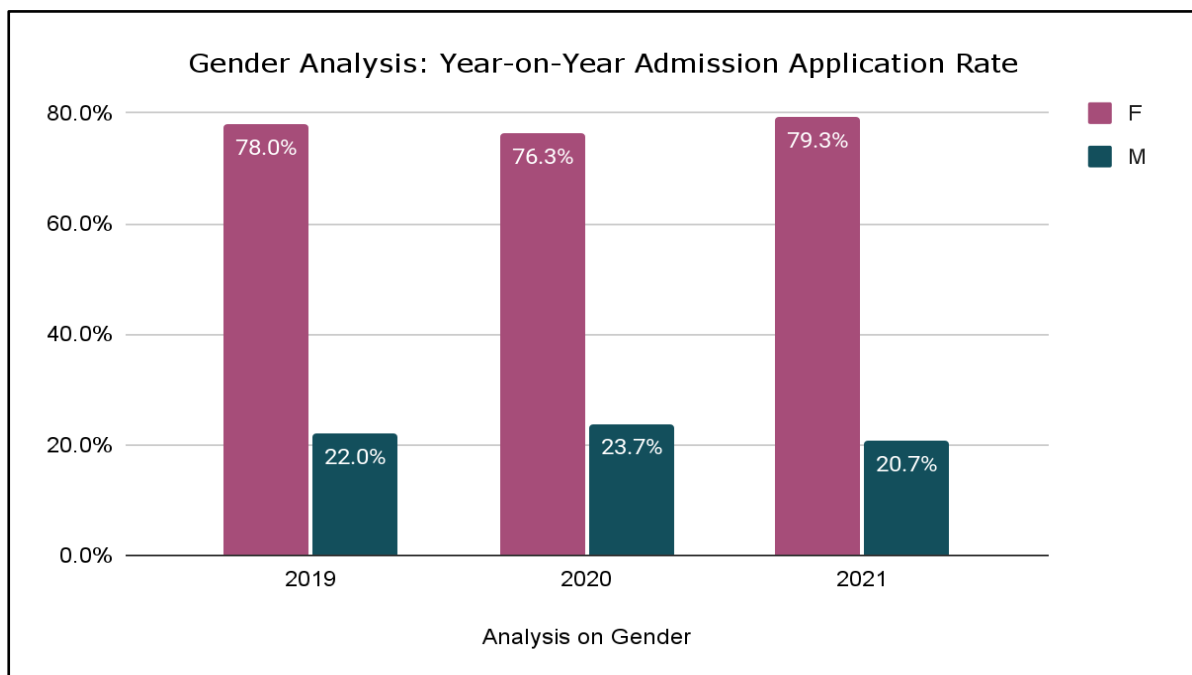


Figure 3: Gender Analysis - Yearly Admission Application Rate

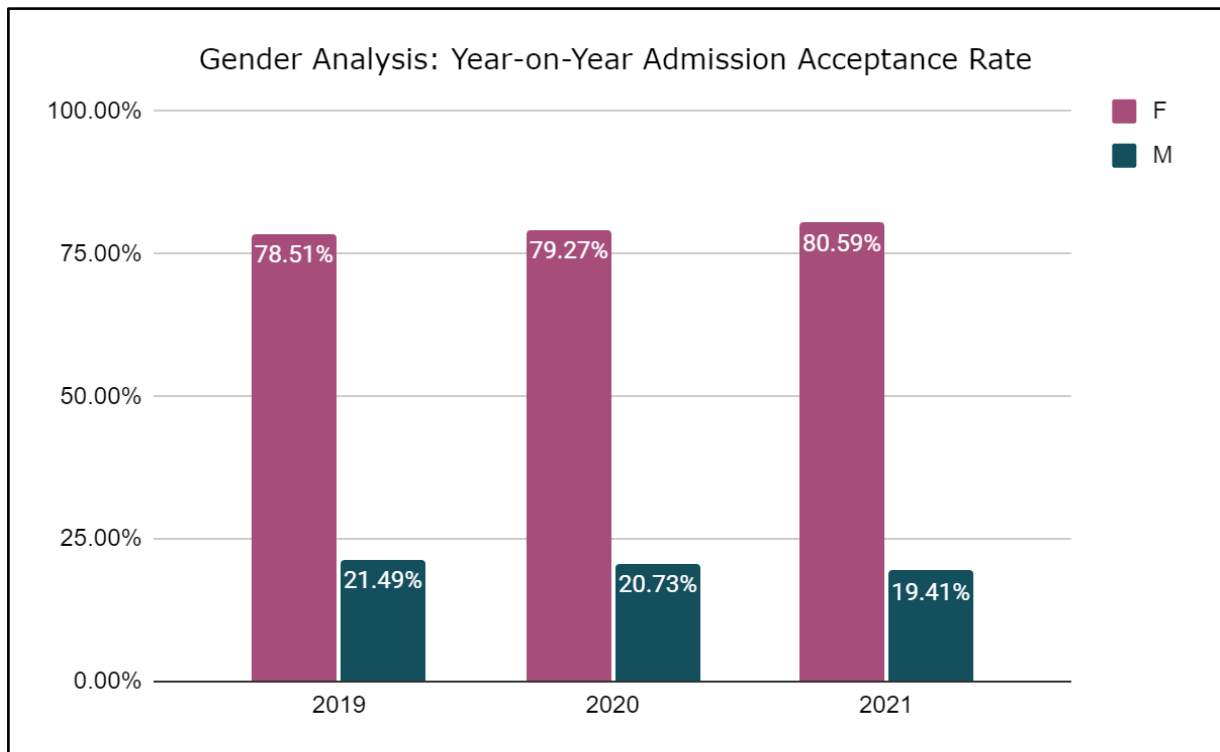


Figure 4: Gender Analysis - Yearly Admission Acceptance rate

3. Geolocation Analysis:

The charts below show the various regions in the world where the university has received applications. All three years have shown that most of the applications came from Botswana. In 2019, a couple of applications were received from other African countries, the United States and India. In 2020, very few applications, less than 1% of applications, were received from a couple of African countries around Botswana and few applications in the United States and India. In the year 2021, more applications were received from African countries than the previous year, 2020. No applications came from the United States and few other applications came from India. This analysis has shown that the university has not excelled at attracting applicants from other parts of the continent or other continents. More marketing needs to be embarked on by the university to showcase what the university has to offer.

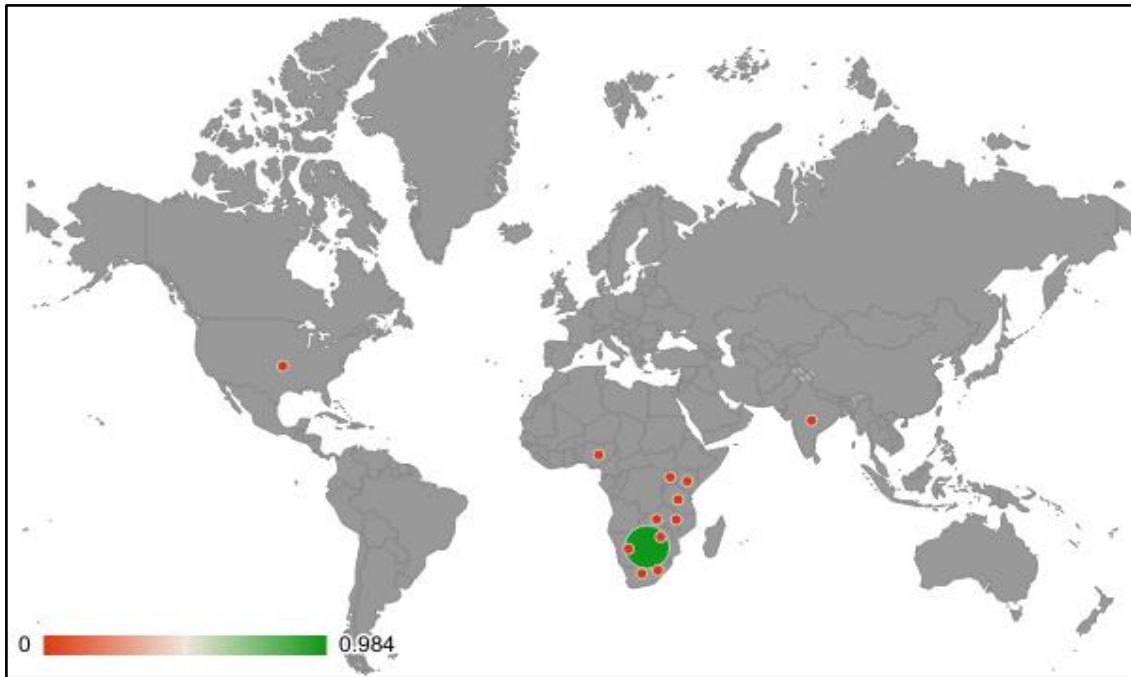


Figure 5: Geolocation Analysis - 2019

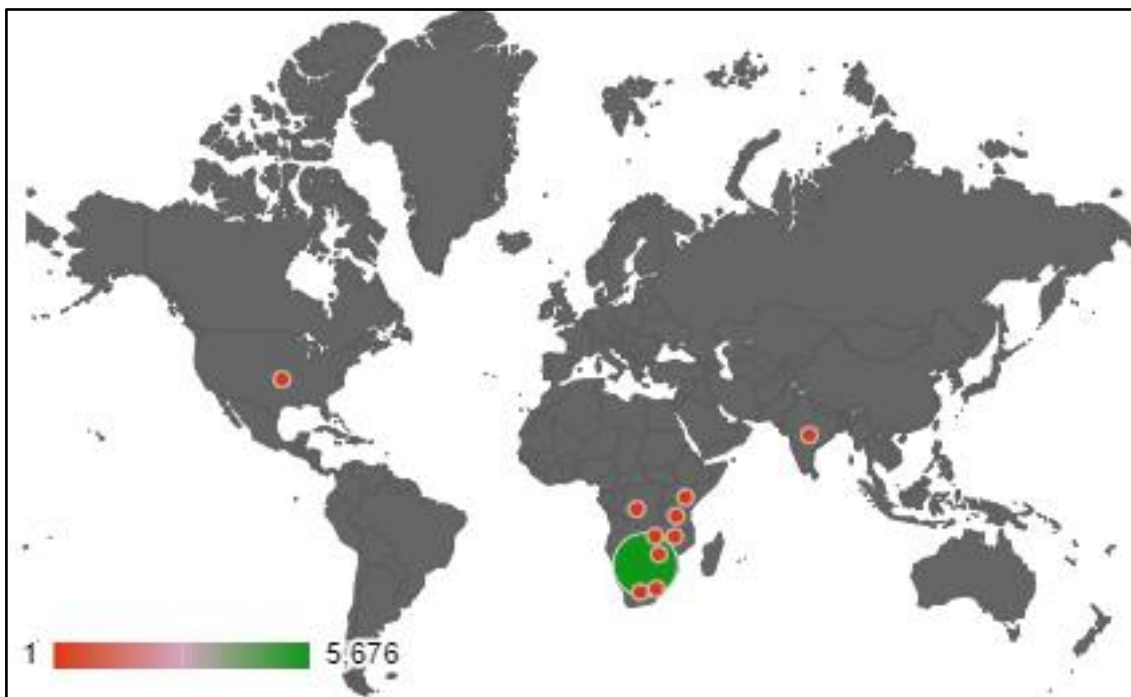


Figure 6: Geolocation Analysis - 2020

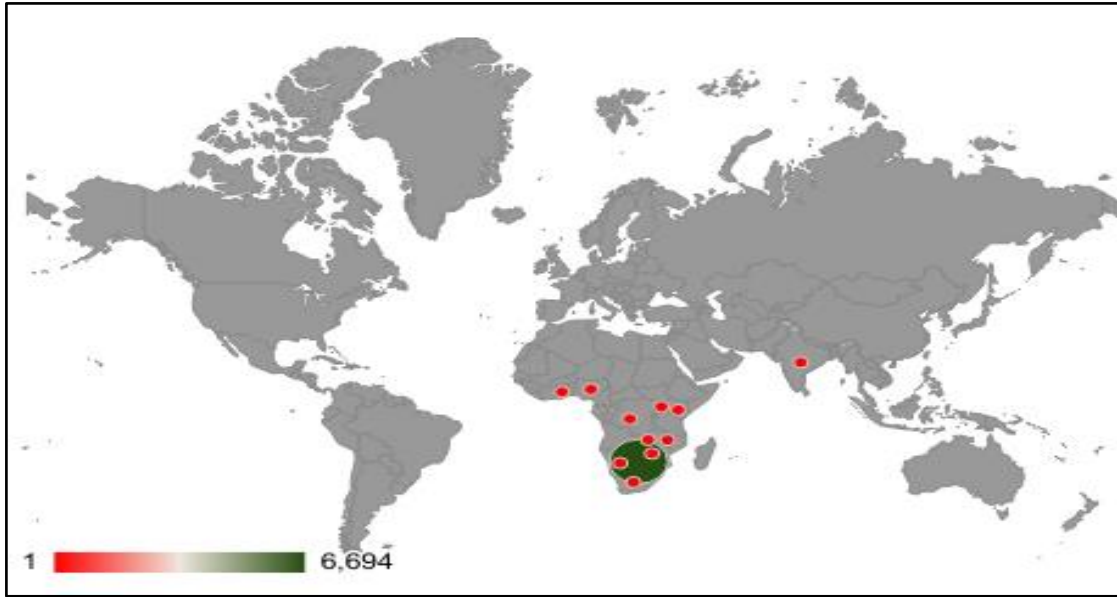


Figure 7: Geolocation Analysis - 2021

	APPLICANTS COUNTRY	2019	2020	2021
1	BOTSWANA	98.38%	97.44%	97.41%
2	ZIMBABWE	1.00%	0.96%	1.14%
3	SOUTH AFRICA	0.10%	0.03%	0.09%
4	OTHER	0.04%	0.05%	0.06%
5	NAMIBIA	0.01%	0.00%	0.01%
6	ZAMBIA	0.14%	0.17%	0.17%
7	LESOTHO	0.01%	0.36%	0.00%
8	KENYA	0.06%	0.09%	0.19%
9	NIGERIA	0.01%	0.00%	0.09%
10	NAN	0.00%	0.69%	0.74%
11	TANZANIA	0.00%	0.05%	0.00%
12	MALAWI	0.08%	0.03%	0.01%
13	INDIA	0.07%	0.05%	0.01%
14	UNITED STATES OF AMERICA	0.00%	0.02%	0.00%
15	UGANDA	0.06%	0.00%	0.01%
16	GHANA	0.01%	0.00%	0.01%
17	DEMOCRATIC REPUBLIC OF CONGO	0.03%	0.05%	0.04%

Figure 8: Cross Tabulation of Applicants Countries of Origin

Due to the fact that the majority of applications were from Botswana, further investigations were done on the Region data. It shows that the 'HQ West Wing Region' accounts for the majority of the applications received by the university, While applicants from the 'Head Quarters' region were the least. Further analysis will be done to understand why the majority of the applications were from the HQ West Wing.

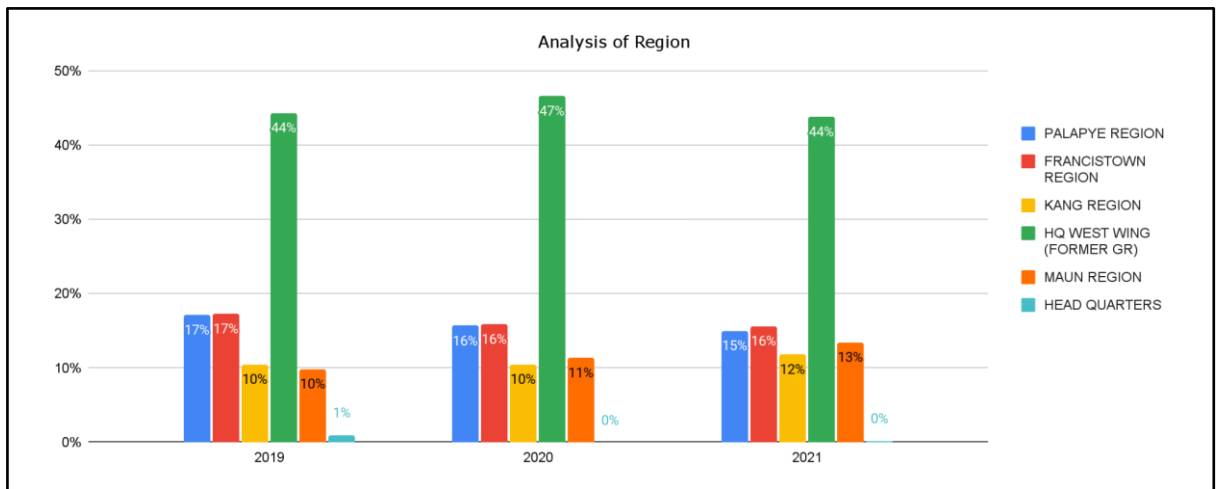


Figure 9: Analysis of Regions in Botswana

4. Program Types Analysis:

Over the years, the rate of applications for a Post Graduate program is slightly greater than the rate of applications for an Undergraduate program. This would imply applicants seek an Undergraduate degree almost at the same rate as a Postgraduate degree in the university.

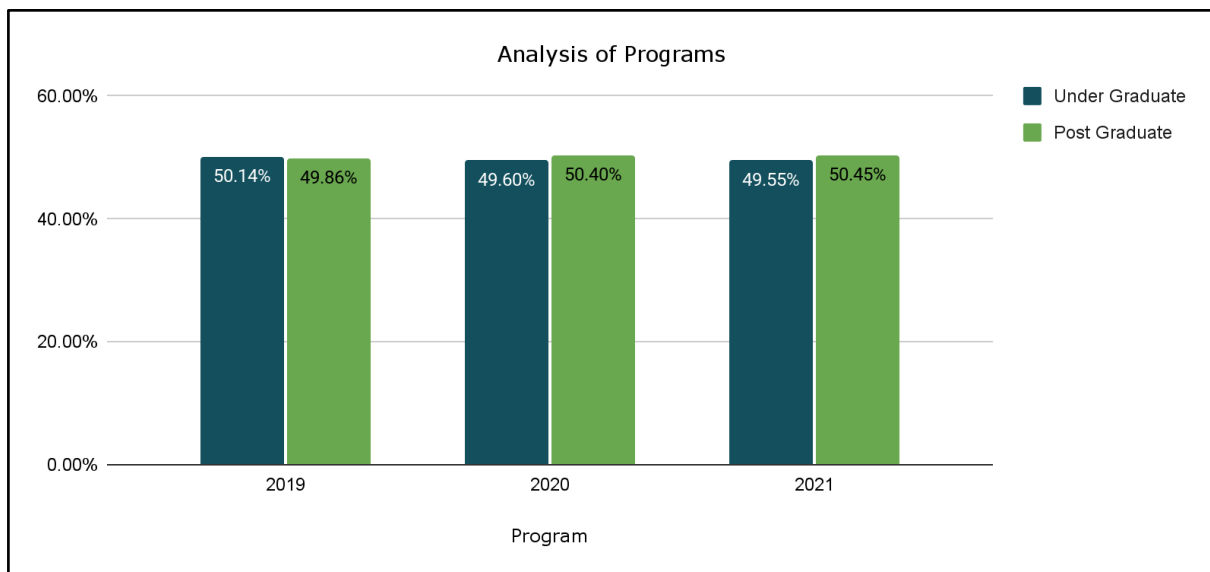


Figure 10: Yearly Analysis of Program Types

After further analysis, it became evident that the undergraduate applicants have a higher rate of acceptance into the institution than the postgraduate applicants. Even during the COVID-19 pandemic, the rate of acceptance for undergraduate applicants was not affected but the rate of acceptance for the postgraduate applicants was affected drastically, with over a 10% drop in the acceptance rate.

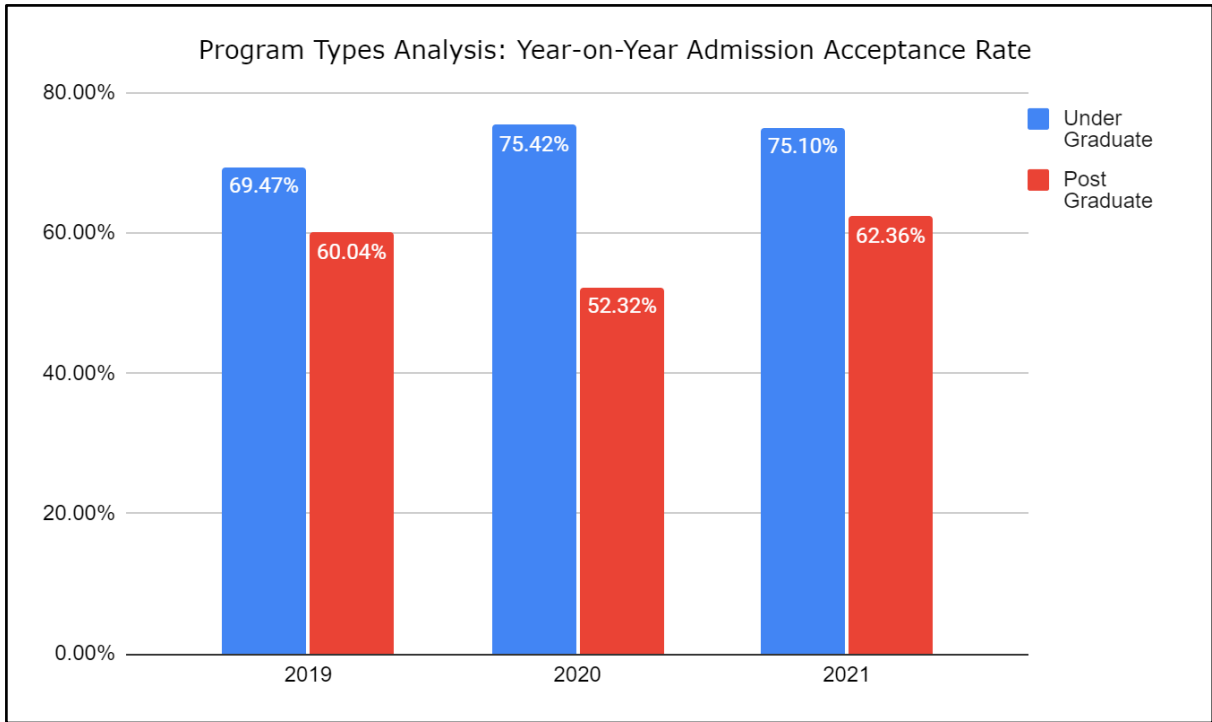


Figure 11: Program Types Analysis - Yearly Admission Acceptance Rate

5. Qualification Analysis:

The majority of the applicants are either Diploma degree holders or possess a Vocational certificate. Very few applicants had a skills certificate or a secondary school application. This shows that individuals with Diplomas and Vocational certificates tend to apply to study at the university. This aligns with what was detailed earlier in the program types of analysis. The low to zero application rate for those with Secondary school Education may be because of the low acceptance of open university study within that circle. Good publicity on the benefits of the open university across all strata should be embarked on.

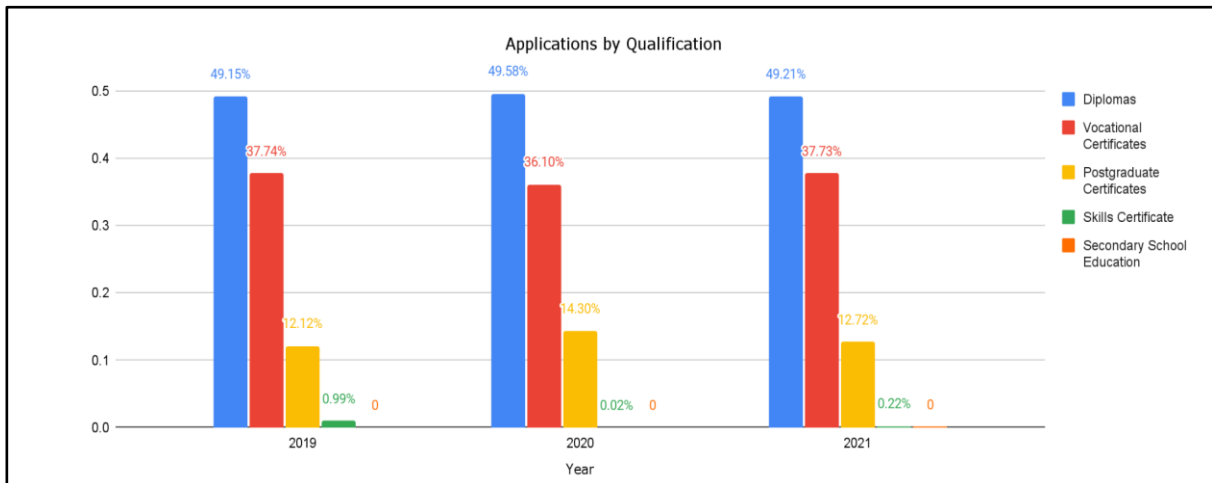


Figure 12: Yearly Applications by Qualifications

6. Students Employers Analysis:

Majority of applicants are employees of the Ministry of Education followed by unemployed applicants. The Botswana Institution of CA had the least number of applicants. The institution could further resources into attracting more of their biggest clients, the Ministry of Education.

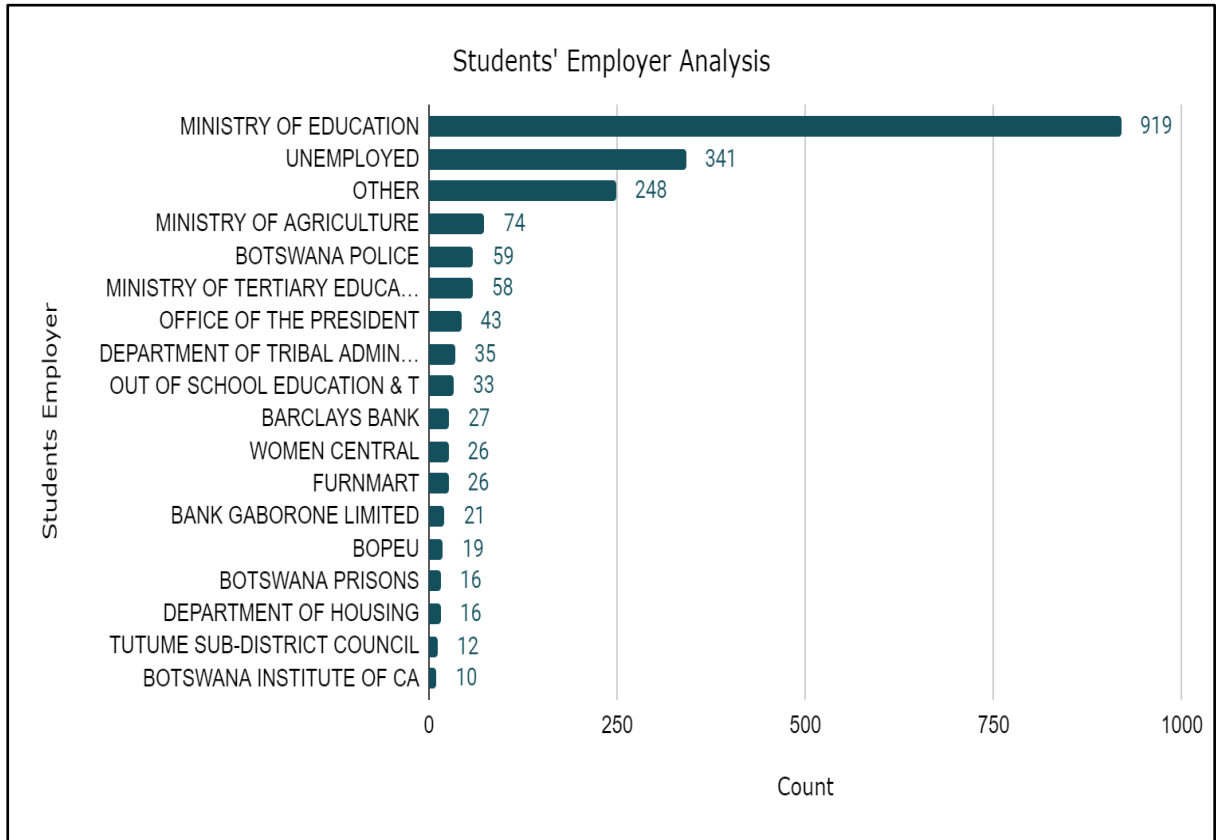


Figure 13: Students' Employer Analysis

CONCLUSION

In order to understand the system from which the data were to be collected, a demo application process was undertaken through which several observations were made; the date of birth is not restrictive to any date/period which made it possible for the applicant to input the date of their application as the date of birth, meaning that some date of birth may not necessarily be a true reflection of the age of applicants. This was further observed on the age of applicants analysis where ages 0 and 121 were recorded. The system is also at risk of SQL injection as the demo applicant was able to tamper with it and change their first name on the URL and the system recorded that particular first name over the one which was initially stated in the application form. There was missing data on some variables such as the decision date which was quite common on the rejected applications; this results in an information gap on when a decision to reject a particular application was made.

RECOMMENDATIONS

Data shows that it is not very common for many fresh high school leavers to join the Open University as most of the applicants are those who have prior qualifications. This is an opportunity for BOU to strengthen their marketing so as to attract high school leavers and encourage them to embrace the culture of self-learning. With the Covid-19 pandemic, traditional universities have designed and implemented robust online learning applications to attract students from all over the world, BOU should leverage this trend, and welcome and advertise the new normal to drive students acquisitions and retention. The university should as a matter of urgency resolve the admission portal lapse that leads to SQL injections via the URL, as this can make the university's application portal a soft and easy target for malicious persons. The system should be revamped, made more intuitive, restricted values that are inline with the University goals, and top online security should be implemented.

Software applications with the ability to track students' movement on the University's website and portals should be implemented. This would help identify where students are dropping off on the site, and help inform the design or purchase of a world-class student portal.

This study sets the tone for data mining, analysis and interpretation at the Botswana Open University and the researchers recommend that further analysis for this particular study should be used to investigate attrition rate in the University and map its financial impact. A machine learning model can be developed to forecast possible applications and enrollments for the next 3 years, or as may be determined by the interested stakeholders.

REFERENCES

- Ali, M. M. (2013). ROLE OF DATA MINING IN EDUCATION SECTOR. *International Journal of Computer Science and Mobile Computing*, Vol. 2(4).
- Aulck, L., Nambi, D., & West, J. (2020). *Increasing Enrollment by Optimizing Scholarship Allocations Using Machine Learning and Genetic Algorithms*. 10.
- Cirelli, J., Konkol, A. M., Aqlan, F., & Nwokeji, J. C. (2018). *Predictive Analytics Models for Student Admission and Enrollment*. 9.
- Delcours, N., & Carmona, J. S. (2019). Enrollment management analytics: A practical framework. *Journal of Applied Research in Higher Education*, 11(4), 910–925. <https://doi.org/10.1108/JARHE-10-2018-0209>
- Muhammad, R., & McManus, K. (2018). Charting the Course: Using Data Analytics to Assist with Strategic Enrollment Management. *Strategic Enrollment Management Quarterly*, 6, 7.
- Mulugeta, M., & Borena, B. (n.d.). *Higher Education Students' Enrolment Forecasting System Using Data Mining Application in Ethiopia*. 2(2), 8.
- Nandeshwar, A. R. (n.d.). *Enrollment Prediction Models Using Data Mining*. 18.
- Sabnani, H., More, M., Kudale, P., & Janrao, S. (n.d.). *Prediction of Student Enrolment Using Data Mining Techniques*. 05(04), 5.
- Shrestha, R. M., Orgun, M. A., & Busch, P. (2016). Offer acceptance prediction of academic placement. *Neural Computing and Applications*, 27(8), 2351–2368. <https://doi.org/10.1007/s00521-015-2085-7>
- Soltys, M., Dang, H., Reilly, G. R., & Soltys, K. (2021). Enrollment Predictions with Machine Learning. *Strategic Enrollment Management Quarterly*, 9, 9.
- Zhang, Q. (2012). Application of Data Mining in Analysis about Source of Students of Civil Universities. *Applied Mechanics and Materials*, 241–244, 2827–2832. <https://doi.org/10.4028/www.scientific.net/AMM.241-244.2827>