

DISTRIBUTION AND PREDICTIVE OWNERSHIP OF PRIMARY SCHOOLS IN TANZANIA MAINLAND

Privatus Christopher Ntare, Assistant Lecturer, Department ICT and Mathematics, College of Business Education (CBE), P.O.BOX 1968, Dar es Salaam; Tanzania E-mail: p.ntare@cbe.ac.tz

ABSTRACT

This study attempted to explore the distribution of primary schools and the predictive ownership of primary school teachers in Tanzania's Mainland. The dataset used in this study was collected in 2016 by the Tanzania National Bureau of Statistics. To attain the objectives of this study the descriptive statistical analysis and logistic regression model were adopted. A total of 17,094 schools were included in the study. The findings revealed that there was an unequal distribution of primary schools as well as primary school teachers in Tanzania's Mainland. Tanga was the region with many government primary schools (992) and Dar es Salaam was the region with many non-government schools (245). Katavi was the region with few government primary schools (175) and non-government schools (2). There were more permanent female teachers (104052) than permanent male teachers (98028) while in the case of the temporary teachers, males (2660) were more than females (1970). Also, the data showed that 94.1% of all schools were owned by the government. The logistic regression results indicated that the permanent female teachers were negatively connected with non-government schools while the permanent and temporary male teachers were likely to be in non-government schools. The logistic regression model successfully predicted the school ownership for 13663 schools with a reasonably high accuracy of 94.75%.

Keywords: Primary Education, Teachers Distribution, School ownership, Logistic regression.

INTRODUCTION

Background Information

Education is the basic tool for the social, political, and economic development of any country. Since the adoption of the Universal Declaration of Human Rights in 1948, education is recognized as a basic human right besides shelter and health. Education for All (EFA) was adopted by more than 150 governments since its establishment in 1990 as the policy to support the universal right to education (Madani, 2019). Since then, various governments have struggled to promote the provision of quality education through the implementation of different education policies and strategies. Like other countries in the world, the Tanzanian government did not run behind the universal commitment of providing quality education for all children, youth, and adults. To promote the EFA policy, it established the education and training policy of 1995 which stipulated the structure of the formal education and training as 2-7-4-2-3+, which means that 2 years of pre-primary education, 7 years of primary education, 4 years of secondary ordinary level education, 2 years of the advanced secondary education and 3 or more years in higher education (Mghasse & William, 2016).

Since the 1960s the developing countries like Tanzania have identified primary education as the priority target (Haulle & Kabege, 2020). The primary education in Tanzania mainland laid on the philosophy of Julius Kambarage Nyerere who was the founder of education for self-reliance (Nyerere, 1967). The major aim of giving primary education to Tanzanians was to provide them with the knowledge for a self-reliant life. This made primary education to be the ultimate education for the majority (Chacha & Zhong, 2013). The rapid expansion of primary education came after the 1974 Universal Primary Education (UPE) programme where many schools were built and many teachers were recruited.

From the 1990s to the present, the government of the United Republic of Tanzania have decided to offer the primary education free to every child. This motivated many parents to take their children to school and the number of students has been increasing each year.

According to Valente, (2019), the net enrollment in Tanzania primary school went from 53% in 2000 to 73% in 2002, coinciding with the removal of primary school fees. Another initiative was the Primary Education Development Programme (PEDP) which was introduced by the government in 2001 with support from international donors (Dennis

&Stahley, 2012). To implement the PEDP, many reforms were introduced and one of them was the double-shift system in public schools. Dar es Salaam region was selected as the pilot region and the findings revealed that there was no strong impact of multiple shift systems (Mbelle, 2008). As an effort to improve education in Tanzania, the education sector was divided into four sub-sectors which included basic education, folk education, higher education, and technical and vocational training education (Lusingu, 2013). Also, the government introduced the 2025 vision which aimed to eliminate poverty by providing better education to all Tanzanians (URT, 1999).

According to Mbelle (2008), the efforts of achieving universal primary education leads to the implementation of initiatives such as increasing the funding levels, provision of textbooks, mobilization of external support from different stakeholders, improving learning and teaching environment by providing the teaching and learning materials and increasing the number of teachers to attain the reasonable teacher-pupil ratio. Several studies have been done to examine the link between policy and practice at the primary school level. For example, Haulle & Kabelege, (2020) examined the relevance and quality of social science textbooks used in primary schools and found that there was no harmony of contents and some textbooks had wrong contents.

Through this background, it was noted that there is little knowledge regarding the distribution of primary schools and primary school teachers in Tanzania mainland and more specifically on how primary school teachers relate to school ownership. Therefore, it was the intention of this study to explore the school distribution and predictive ownership of primary school teachers to provide valuable insight into how educational stakeholders may be able to improve the provision of education in Tanzania mainland.

METHODOLOGY

Study Area

This study was conducted in Tanzania mainland which formerly was known as Tanganyika which became a sovereign state on 9th December, 1961. It is located in Eastern Africa between the longitude 29^oEast and 40^oEast and latitude 1^o to 12^o in the Southern hemisphere. Currently, Tanzania mainland has 26 regions which include Arusha, Dar es Salaam, Dodoma, Geita, Iringa, Kagera, Katavi, Kigoma, Kilimanjaro, Lindi, Manyara, Mara, Mbeya, Morogoro, Mtwara, Mwanza, Njombe, Pwani, Rukwa, Ruvuma, Shinyanga, Simiyu, Singida, Songwe, Tabora, and Tanga. Tanzania mainland covers about 942,800km² and its population in 2012 was 43,625,354 (Agwanda & Amani, 2014).

Data Collection and Preparation

The dataset used in this study was collected in 2016 by the National Bureau of Statistics (NBS). The data comprises 9 variables which include the regional, council, ward, school name, and school ownership, permanent male, permanent female, temporary male, and temporary female teachers in each school. The data was collected annually from all 26 regions during the basic education census. The President's Office Regional and Local Government (PO-RALG) sent data collection forms to be filled by primary school headteachers in each region. All headteachers filled in the forms and sent them to the Ward Education Coordinator's office (WEC) for verification and certification of the information filled in the forms. After verification, the WECs sent the forms to District Executive Director (DED) office where the data was entered into the Basic Education Information System. The system used is web-based with a direct link to the PO-RALG database. This study used the data as they were presented by NBS.

Logistic Regression for Machine learning

Logistic regression is one of the techniques borrowed by machine learning from the field of statistics. It is used in modeling the data when the response variable is in a categorical form. Binary logistic regression is the special type of regression where the response variable is related to the set of explanatory variables that can be discrete or/and continuous (Agresti, & Kateri, 2011). According to Tibshirani *et al* (2015), if the binary response coded in the form of $Y \in \{0,1\}$ the linear logistic model is often used and it models the log-likelihood ratio as the linear combination

$$\log \left(\frac{P(Y=1|X=x)}{P(Y=0|X=x)} \right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \dots \dots \dots (1)$$

Where $X = (X_1, X_2, \dots, X_p)$ are p explanatory variables, $\beta_0 \in R$ is an intercept term and $\beta_1, \dots, \beta_p \in R$ are the vector regression coefficients. Inverting this transformation yields an expression for conditional probability

$$P(Y = 1|X = x) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)} \dots \dots \dots (2)$$

By inspection without any restriction on the parameter $(\beta_0, \dots, \beta_p)$, the model specifies probability lying in $[0,1]$. Typically logistic models are fitted by maximizing the binomial log-likelihood of the data.

Prepare data for Logistic Regression

In this study, the aim was to create a predictive machine learning model that can accurately predict than interpreting the results. The model was built to predict the probability of a teacher to belong in the default class which can be either government or non-government. The outlier from the training data was checked to avoid the model overfitting, then the pairwise correlation between all explanatory variables was calculated and highly correlated variables were removed.

Learning the Logistic Regression Model

The coefficient (β) of the logistic regression, the algorithm was estimated from the training data by using the maximum likelihood estimation. The best coefficient resulted in the model that predicts the value greater or equal to 0.5 as 1 (government) for defaults class and value less than 0.5 close to 0. The intuition of the maximum likelihood for logistic regression was to look at coefficients that minimize the error in the predicted probability model to those in the data.

Training the logistic Model and Making a prediction

The model was trained on two completely separate datasets. 80% of the observations on the data were randomly selected as the training dataset and testing was performed using the rest 20% of the observations. Making the prediction we plugged the number into the logistic regression equation and calculated the result. Finally, we computed the prediction of 20% observations and compared the results with the actual observations.

Confusion Matrix

A confusion matrix shows the number of correct and incorrect predictions made by the classification model compared to the actual outcomes in the data (Tibshirani *et al*, 2015). The output of the confusion matrix tells us how the model is performing. It is summarized in Table 1.

Table 1: The confusion matrix describing the agreement between the actual and predicted values.

CONFUSION MATRIX			Actual			
			Positive	Negative		
Predicted	Model	Positive	A	b	Positive predicted value	a/(a+b)
		Negative	C	d	Negative predicted value	d/(c+d)
			Sensitivity	Specificity	Accuracy = (a+d)/(a+b+c+d)	
			a/(a+c)	d/(b+d)		

RESULTS AND DISCUSSION

In this study, a total of 26 regions, 180 councils, 3549 wards, and 17094 schools were recorded. The results of this section were obtained using the statistical software environment R.

The Distribution of Primary school teachers in Tanzania.

From the dataset, the distribution of primary school teachers based on the categories is shown in Table 2.

Table 2: The Distribution of Primary School Teachers Based on Their Categories and Sex.

		Ownership			
Sex	Teachers	Government	Non-Government	Total	
		Permanent Male	91630	6398	98028
	Temporary Male	466	2194	2660	
	Permanent Female	99140	4916	104056	106026
	Temporary Female	510	1460	1970	
Total		191746	14968	206714	

Source: Field Data.

From Table 2, we observe that the Tanzania mainland had a total of 206714 primary school teachers whereby 191746 were in government schools and 14968 were in non-government schools. Permanent male teachers were 98028, permanent female were 104056, temporary male were 2660 and temporary female were 1970. We also observed that permanent teachers (202084) are more than temporary teachers (4630), and female teachers (106026) are more than male teachers (100688). There were more permanent male teachers (91630) employed in government schools than temporary male teachers (466), and permanent female teachers (99140) were employed in government schools than temporary female teachers (510). From these results, we concluded that there was no gender balance in the distribution of primary school teachers in Tanzania mainland.

The Regional Distribution of Primary Schools in Tanzania

The percentage distribution of primary schools in 26 regions of Tanzania Mainland was summarised by a pie chart in Figure 1.

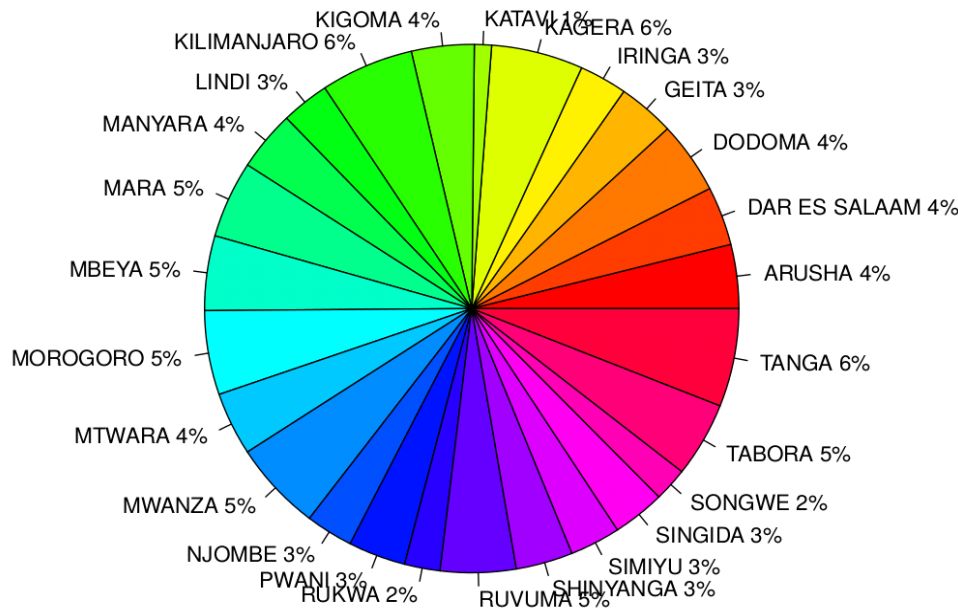


Figure 1: Pie Chart to Show the Percentage Number of Primary School Teachers per Region

Figure 1 shows that there is an unequal distribution of primary school teachers in Tanzania mainland. Tanga, Kagera, and Kilimanjaro regions had a high number of primary school teachers (6%) while Rukwa (2%), Songwe (2%), and

Katavi (1%) had the lowest number of teachers. The low number of teachers in the regions with many rural areas such as Rukwa, Simiyu, Katavi, and Songwe may be the pushing factors that make teachers to get out of their working stations due to the unreliable working environment that leads to inequitable distribution of teachers (Lusingu, 2013).

The Regional Distribution of Primary School Teachers by Sex

The regional distribution of primary school teachers by sex in Tanzania Mainland was summarised by using histogram in Figure 2.

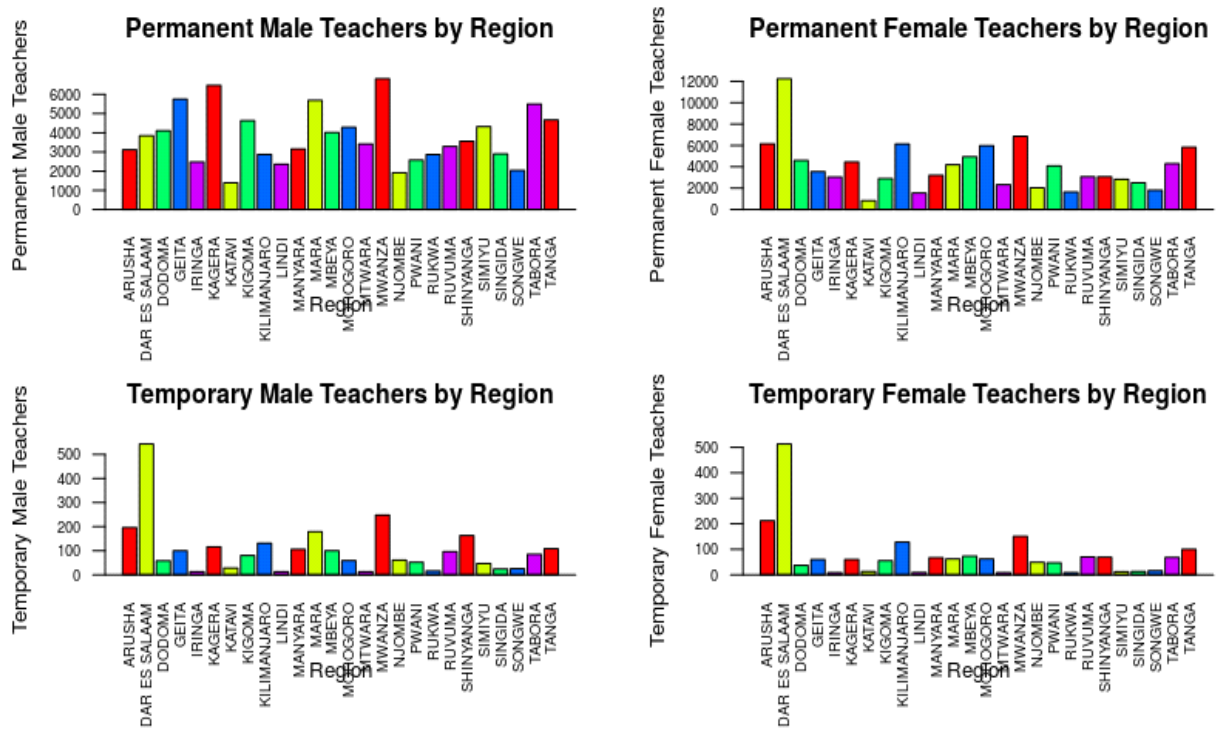


Figure 2: Regional Distribution of Primary School Teachers by Sex in Tanzania

From Figure 2 it is observed that Mwanza, Kagera, Geita, Mara, and Tabora region have many permanent male teachers while Katavi, Njombe, and Songwe have a low number of permanent male teachers. Also, in the case of permanent female teachers, Dar es Salaam is the leading region followed by Mwanza, Arusha, and Tanga while Lindi, Rukwa, and Njombe have a low number of permanent female teachers.

For the case of temporary male teachers, Dar es Salaam is leading followed by Mwanza, Arusha, Shinyanga, and Tanga while Iringa, Lindi, Mtwara, Rukwa, Singida, and Songwe have few temporary male teachers. Also, we observe that Dar es Salaam has the highest number of temporary female teachers followed by Arusha, Mwanza, and Tanga. On the other hand, Mtwara, Rukwa, Simiyu, Singida, and Songwe region have a lower number of temporary female teachers.

This indicates that in Tanzania mainland there were many permanent and temporary primary teachers of both sexes in the regions that have cities like Dar es Salaam, Mwanza, Arusha, and Tanga than the regions with no city or big towns like Simiyu, Songwe, and Katavi. This may be due to the presence of good infrastructures such as teachers' houses, water, electricity, roads, and other social services in the region with cities which motivate teachers to stay in their workstations than the regions with many rural areas where these facilities are limited.

The Regional Distribution of Primary Schools by Ownership

The regional distribution of primary school teachers by ownership (government or non-government) is summarised in Figure 3.

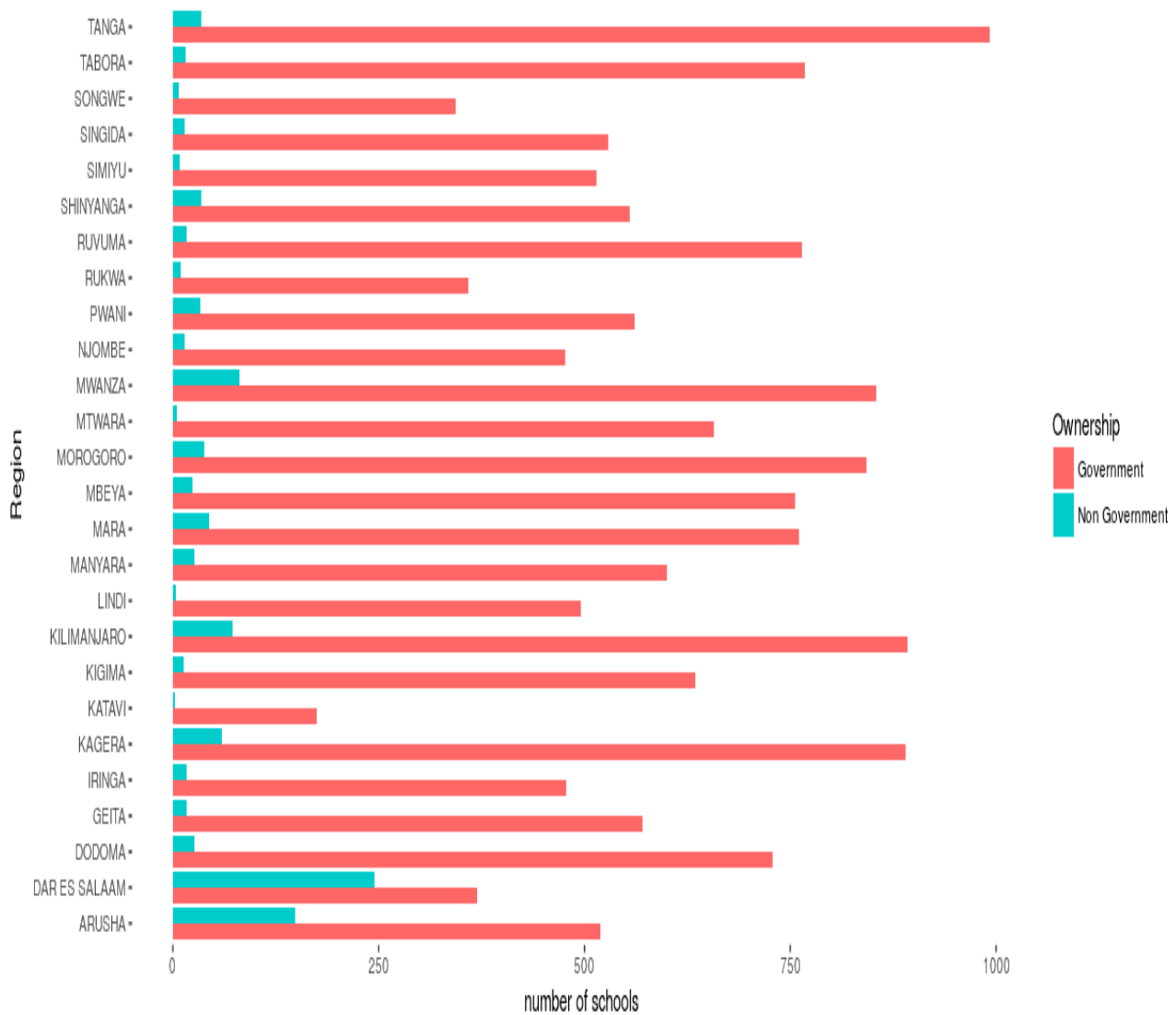


Figure 3: Regional Distribution of Primary School Teachers by Ownership

From Figure 3 it is observed that there were more government schools than non-governmental schools in all regions. Also, it was observed that Tanga was the region with many governmental primary schools (992), and Dar es Salaam was the region with many non-governmental schools (245). Katavi was the region with few governmental primary schools (175) and non-governmental schools (2). This may be due to the reason that both are city regions with a high population, good social services as well as good infrastructures than other regions. The regions with cities like Dar es Salaam, Arusha, Tanga, Mwanza, and Kilimanjaro have undergone major economic changes since independence, they have educational resources and infrastructures that are required for the provision of quality education to students. Compared to the regions with the developed township, the regions with many villages like Songwe, Rukwa, and Katavi have poorly developed education infrastructures and limited social services which leads to a few primary schools and teachers.

The Compositional Relation between Primary school Teachers

To get more insight into the relationship between different categories of primary school teachers, we used the correlation matrix and the results are indicated in Figure 4.



Figure 4: Variables Correlation Matrix.

The compositional relation in Figure 4 shows that there is a high positive linear correlation (0.67) between temporary male and temporary female teachers, this indicates that these variables vary together. This implies that the school with a high number of temporary male teachers had a high number of temporary female teachers and vice versa was true. Also, we observed that there was a very weak negative relation between temporary female and permanent male teachers (-0.11) and between temporary male and permanent male teachers (-0.13). The correlation between permanent male and permanent female teachers was positive but very weak (0.07). This indicated that these variables did not depend on one another.

The Relationship between School Ownership and Teachers

From the dataset, school ownership was a dichotomous (binary) variable with two levels (government and non-government), we used logistic regression to model the school ownership variable. In the logistic model, the log odds of the outcomes were modeled as the linear combination of the explanatory variables. We let the value of 1 if the school owner was the non-government and 0 if the owner was the government. The Permanent Male (PM), Permanent Female (PF), Temporary Male (TM), and Temporary Female (TF) were the explanatory variables that may influence the school ownership (Response variable) probability. The initial model where $P = Pr(owner = 1|PM, PF, TM, TF)$ were considered.

For this model deviance $2(l_s - l_m) = 4794.6$ with 13659 degrees of freedom and p-value equal to 1, whereby l_s denoted the logarithm of the likelihood function for the saturated model and l_m was the logarithm of the likelihood function for the proposed model.

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 PM + \beta_2 PF + \beta_3 TM + \beta_4 TF + \varepsilon \dots \dots \dots (3).$$

The binomial distribution with the logit link function was used to run the analysis. After fitting the model using the training dataset, TF was eliminated since it has a higher p-value and we obtain the fitted model results as summarized in Table 3.

Table 3: The Logistic Regression Results

Coefficient	Estimate	95% Confidence Interval	Std. Error	z -value	p-value
Intercept	-3.8798	(-4.0453, -3.7177)	0.0835	-46.441	0.0000
Permanent Male (PM)	0.1316	(0.1126, 0.1506)	0.0097	13.599	0.0000
Permanent Female (PF)	-0.0152	(-0.0276, -0.0036)	0.0061	-2.485	0.0129
Temporary Male (TM)	0.9946	(0.9106, 1.0843)	0.0443	22.469	0.0000

The fitted model has the deviance of 6133.2 with 17090 degrees of freedom and a p-value equal to one. From Table 3 the fitted model was

$$\text{logit}(\hat{p}) = -3.8798 + 0.1316PM - 0.015PF + 0.994TM \dots \dots \dots (4)$$

The logistic regression coefficients give the change in the log-odds (logit \hat{p}) of the school ownership for a unit change in the predictor variable. Equation 4 indicates that if the other variables are kept constant then for each unit increase in PM, the log odds of non-governmental school increases by 0.131. For every unit change in TM, the log odds of non-government schools increase by 0.99. And for every unit change in PF, the log odds of non-government school decreases by 0.015. The goodness of fit test using the chi-square test shows that the p-value was $9.487085 \times 10^{-279}$ which indicates that the model was statistically significant. Therefore, we conclude that if all other variables were constant, the permanent female teacher was negatively related to non-governmental schools while the permanent and temporary male teachers were likely to be related to non-governmental schools. To predict using logistic regression, we used the confusion matrix and the results are summarised in Table 3.

Table 4: Confusion Matrix and statistics for the training dataset.**Confusion Matrix and Statistics**

Prediction	Reference	
	Government	Non-Government
Government	12820	606
Non-Government	45	192

Accuracy : 0.9524
 95% CI : (0.9486, 0.9559)
 No Information Rate : 0.9416
 P-Value [Acc > NIR] : 1.84e-08

 Kappa : 0.3537
 McNemar's Test P-Value : < 2.2e-16

 Sensitivity : 0.9965
 Specificity : 0.2406
 Pos Pred Value : 0.9549
 Neg Pred Value : 0.8101
 Prevalence : 0.9416
 Detection Rate : 0.9383
 Detection Prevalence : 0.9827
 Balanced Accuracy : 0.6186

 'Positive' Class : Government

From Table 4, the diagonal elements of the confusion matrix represent correct prediction, while the off-diagonal represents incorrect prediction. The model correctly predicted that 192 were non-governmental schools and 12820 were governmental schools. This gives a total of 13012 correct predictions with an accuracy of 0.9524. In this case, the logistic regression correctly predicted the ownership of the school for an accuracy of 95.2%.

At this stage, it appears that the logistic regression model is working better than guessing but this result is misleading because we trained and tested the model on the same set of 13663 observations (training dataset). In other words, the misclassification error in the training dataset is 4.8%. For a better assessment of the model accuracy, we used the held out (test) data to examine how well the model predicts since the training error tends to underestimate the test error rate (James *et al*, 2015).

Results for the confusion matrix of the test data are summarised in Table 4.

Table 5: Confusion Matrix and Statistics for the Test dataset

```
-----
Confusion Matrix and Statistics

          Reference
Prediction Government Non-Government
Government      3204      165
Non-Government    15       47

          Accuracy : 0.9475
          95% CI   : (0.9395, 0.9548)
    No Information Rate : 0.9382
    P-Value [Acc > NIR] : 0.01136

          Kappa   : 0.3242
  Mcnemar's Test P-Value : < 2e-16

          Sensitivity : 0.9953
          Specificity : 0.2217
    Pos Pred Value   : 0.9510
    Neg Pred Value   : 0.7581
          Prevalence : 0.9382
    Detection Rate   : 0.9338
  Detection Prevalence : 0.9819
    Balanced Accuracy : 0.6085

'Positive' Class : Government
```

From Table 5 results, the accuracy of the model appears to be 0.9475. This indicates that 94.75% of school ownership has been correctly predicted. In other words, the misclassification error in the test dataset is 5.25%. However, the confusion matrix shows that the negative and positive predicted values are 0.7581 and 0.9510 respectively. This is an indication that when the logistic regression predicts the non-governmental school has a 75.8% accuracy rate and when it predicts the government school, it has a 95.1% accuracy rate. This shows that the model performs better when it predicts government school ownership rather than non-governmental ownership. By comparing the accuracy of the train and test dataset, one can observe a small decrease in model accuracy. In a further study, one would need to investigate more carefully whether the small decrease in the accuracy of the model was real or just due to random chance.

Finally, we used the chi-square test to perform the goodness of fit test, the model p-value was $9.487085 \times 10^{-279}$ which is too small. Therefore, the confidence level is very high which indicates that the logistic model used was statistically significant.

CONCLUSION

Despite the different efforts taken by the government to improve primary education in Tanzania mainland, the unequal distribution of primary schools and teachers seems to be a problem that hinders the provision of quality education to primary school students. We confirm that there is a role that the government and the non-governmental education stakeholders need to play in order to solve the problem of unequal distribution of primary schools and teachers to catalyze the country's development. The government should play a central role in ensuring the equal distribution of teachers in all regions since it owns many schools (94.1%) and has employed more teachers than the private sector. Then, the government has to introduce and support the major primary education projects in the regions with many rural areas. There is little availability of primary schools and qualified primary school teachers in rural areas. Teachers who are employed in rural areas, whenever they get opportunities, they shift to town and cities. The study therefore recommends that to check this, the government should employ new teachers and appoint other teachers, and send them to the rural primary schools. The strategies such as targeted recruitment, provision of hardship allowances, and incentives to primary school teachers working in rural areas can also be used by the government to make rural posting preferable. Also, the private and public sectors should make up for the joint effort to safeguard equality in the provision of primary education, which has to be guaranteed to all citizens. To conclude, the government should ensure that all citizens are getting a better and quality education. Regions with few primary schools and teachers must be given priorities in education programs to make sure that they do not remain outside the primary education circle in the country.

REFERENCES:

- Agresti, A., & Kateri, M. (2011). Categorical data analysis. In *International encyclopedia of statistical science* (pp. 206-208). Springer, Berlin, Heidelberg.
- Tibshirani, R., Wainwright, M., & Hastie, T. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC Press. New York.
- Agwanda, A., & Amani, H. (2014). *Population growth, structure, and momentum in Tanzania*. Economic and Social Research Foundation.
- Chacha, G., Zhong, Y. (2013). The challenges of Primary Education Level in Tanzania. Case Study Tarime District. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 16 (3), 1-6.
- Dennis, C., Stahley, K. (2012). Universal primary education in Tanzania: The role of school expenses and opportunity cost. *Evans School Review*, 2(1), 47-65.
- Haulle, E., & Kabelege, E. (2020). Relevance and Quality of Textbooks Used in Primary Education in Tanzania: A Case of Social Studies Textbooks. *Contemporary Education Dialogue*, 1-17. <https://doi.org/10.1177/0973184920962702>
- Madani, R. A. (2019). Analysis of Educational Quality, a Goal of Education for All Policy. *Higher Education Studies*, 9(1), 100. <https://doi.org/10.5539/hes.v9n1p100>
- Mghasse, N. E., & William, F. (2016). Practices and Challenges in the Provision of Pre-Primary Education in Tanzania. *African Research Review*, 10(1), 1. <https://doi.org/10.4314/afrev.v10i1.1>
- Valente, C. (2019). Primary education expansion and quality of schooling. *Economics of Education Review*, 73(9208). <https://doi.org/10.1016/j.econedurev.2019.101913>
- Lusingu, J. (2013). Factors for unequal distribution of primary school teachers in Ruangwa District council, Tanzania. *Journal of Education and Practice*, 26 (4), 98-106.
- Madani, R. A. (2019). Analysis of Educational Quality, a Goal of Education for All Policy. *Higher Education Studies*, 9(1), 100. <https://doi.org/10.5539/hes.v9n1p100>
- Mbelle, A.V (2008). *The impact of reform on the quality of primary education in Tanzania*. Retrieved at <https://www.africaportal.org/publications/the-impact-of-reforms-on-the-quality-of-primary-education-in-tanzania/> on 2 June 2018.
- Mghasse, N. E., & William, F. (2016). Practices and Challenges in the Provision of Pre-Primary Education in Tanzania. *African Research Review*, 10(1), 1. <https://doi.org/10.4314/afrev.v10i1.1>
- Nyerere, J.K. (1967). Education for self-reliance. *The Ecumenical Review*, 19(4), 382-403.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2015). *An Introduction to Statistical Learning with Application in R*. New York. Springer.
- President's Office, Regional Administration, and Local Government (PORALG) (2017). *Number of Teachers in Primary Schools at the School level*. Retrieved at <http://statistics.go.tz/dataset/9fa81051-3b66-410a-b461-823942611b10/resource/a20c9afc-c717-4c25-829c18db1f27a642/download/Number-of-Teachers-in-Primary-Schools-2016.csv> on 6 April 2017.
- Sabates, R., Westbrook, J, and Hernandez-Fernandez, J. (2012). The 1977 Universal Primary Education in Tanzania: a historical base for quantitative inquiry. *International Journal of Research & Method in Education*, 35(1), 55-70.
- United Republic of Tanzania (URT). (2004). *Primary Education Development Plan: Progress Report for the financial year 2002/03*. Dar es Salaam, MoEC& PORALG.
- United Republic of Tanzania (URT). (1999). *Development Vision 2025*. Retrieved at www.tanzania.go.tz/vision.htm on 6 May 2018.
- United Republic of Tanzania President's Office Regional Administration and Local Government (URT- PORALG). (2016). *Pre-Primary, Primary, and Secondary Education Statistics in Brief*. Retrieved at <http://www.tamisemi.go.tz/noticeboard/tangazo-1062-20170113-BEST-Regional-and-Pocket-Data-2016/BEST-2016-Pocket-Size-Final.pdf>. On 6 April 2018.
- Valente, C. (2019). Primary education expansion and quality of schooling. *Economics of Education Review*, 73(9208). <https://doi.org/10.1016/j.econedurev.2019.101913>