VOL. 08 NO. 9,
MAY, 2025
E-ISSN 3026-958X
P-ISSN 3027-1169

*African Scholar Publications & Research International*
www.africanscholarpub.com

*Journal of Science Innovation & Technology Research (JSITR)*

# Applying Ordinal Logistic Regression to Analyze Student Performance: A Case Study at Federal Polytechnic Bali

***Mukhtar Sunusi Hassan; & **Usman Abdullahi Sani**

*General Studies Department, School of Business and Management Technology, Federal Polytechnic, Bali. **Department of Statistics, School of science and Technology, Federal Polytechnic Bali, Taraba State – Nigeria

**Corresponding Author:** mukhtarhassansunusi@gmail.com
**DOI:** https://doi.org/10.70382/ajsitr.v8i9.032

## *Abstract*

The persistent decline in the academic performance of Nigerian students in recent years has become a matter of significant concern. Despite the critical importance placed on educational achievement, various studies have documented this downward trend and identified multiple factors that may influence student outcomes. This research utilizes data collected from the records of graduated students at the School of Science, Federal Polytechnic, Bali, covering the academic sessions of 2021/2022, 2022/2023, and 2023/2024. The study investigates the impact of several variables—mode of entry, age at entry, department, and sex—on students' academic performance. Employing ordinal logistic regression (proportional odds model), the analysis reveals that gender does not significantly determine the final academic grade attained at graduation. Notably, both male and female students have an equal likelihood of graduating with first-class honors, suggesting that educational policies should support both genders equally rather than prioritizing one over the other. The findings also indicate that younger students tend to perform better than their older counterparts, implying that age at entry should not be a primary concern; instead, emphasis should be on students' ability to meet academic demands. Furthermore, students with a strong secondary school background, particularly in core science subjects, are more likely to graduate with top honors. Such students are often more academically mature, many having attended private secondary schools with solid

foundational education. Based on these insights, it is recommended that policies promoting programs like the Higher School Certificate (HSC) or Remedial Science be reintroduced to ensure that only academically prepared and mature students gain admission into tertiary institutions. These measures could help improve overall student performance and produce more capable graduates in Nigerian polytechnics.

**Keywords:** Ordinal Logistic Regression, Odds Ratio, Link Function, Students' Performance

# Introduction

Academic performance remains a pivotal concern within educational research, particularly as recent studies indicate a troubling decline in undergraduate achievement, especially within polytechnic institutions (Ugoji, 2008; Egbule, 2004). The importance of academic success is deeply embedded in societal development narratives, especially in Nigeria, where education is regarded as the cornerstone for national progress. The Nigerian society places immense emphasis on education, believing that the active engagement of students in academic activities directly correlates with technological advancement and socioeconomic development. However, despite these laudable aspirations, empirical evidence suggests a persistent decline in student performance across various educational levels, raising concerns about the underlying factors influencing academic outcomes (Ukpong, 2007).

Understanding the determinants of student performance is essential for developing targeted interventions and policies to improve educational standards. Academic performance can be conceptualized as an outcome influenced by a myriad of factors, including demographic variables such as age, gender, financial status, and mode of entry into the institution. Researchers have employed various statistical techniques to analyze these influences, with the goal of identifying significant predictors of academic success or failure (Ebenuwa-Okoh, 2010). For instance, Ebenuwa-Okoh's (2010) study, which utilized simple random sampling to analyze 375 undergraduates, found that factors like age, gender, and financial status did not significantly predict academic performance, suggesting that other variables might be at play.

The complex nature of learning and performance necessitates sophisticated analytical tools capable of capturing the ordinal and categorical characteristics of academic data. Traditional linear regression models may fall short when the dependent variable—such as academic performance—is inherently ordinal, categorized into levels like high, medium, and low achievement. This is where ordinal

logistic regression becomes particularly valuable. It allows researchers to model the relationship between multiple independent variables and an ordinal dependent variable, providing nuanced insights into how different factors influence student performance levels (Agresti, 1997; McCullagh, 1980).

Applying ordinal logistic regression in educational research has gained traction over recent decades, as it offers a robust framework for analyzing categorical outcome variables while accommodating the ordinal nature of performance data (Petersen & Svanes, 1994; Long & Freese, 2006). This approach is especially pertinent in contexts where performance is measured through grades or performance levels, which naturally follow an ordered scale. By leveraging this statistical method, researchers can identify the relative importance of factors such as mode of entry, age, and sex, which may differentially influence performance levels.

This case study focuses on the application of ordinal logistic regression to analyze student performance at the Federal Polytechnic Bali using data from recent academic sessions (2021/2022 to 2023/2024). The study aims to identify significant predictors among variables like mode of entry, age at entry, and gender, thereby providing actionable insights for policymakers and educational administrators. While it is understood that numerous factors influence academic success, this research emphasizes the practical utility of categorical and ordinal data analysis methods in educational settings. Ultimately, the findings are expected to contribute to the ongoing discourse on improving academic outcomes and fostering an environment conducive to student success in Nigerian polytechnics.

In sum, the use of advanced statistical techniques like ordinal logistic regression offers a powerful means to unravel the intricate web of factors affecting student performance. As educational challenges persist, especially in polytechnic institutions, such analytical approaches are instrumental in informing evidence-based strategies that can enhance learning outcomes and support national development goals (Akaike, 1993; McCullagh, 1980; Long & Freese, 2006; 2024).

**Logistic Regression Model**

The development of logistic regression models is an ongoing process that involves estimating regression coefficients to predict the likelihood of a specific outcome (Hosmer, Lemeshow, & Sturdivant, 2013). At its core, a binary logistic regression model can be expressed through various formulations, offering flexibility in interpretation based on the context of application. The most common form is the linear model of the log-odds, or logit, The general linear regression model (LRM)which is defined as:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \log it(\pi_i) = \alpha_{i0} + \alpha_{i1}x_{i1} + \alpha_{i2}x_{i2} + \alpha_{i3}x_{i3} + \ldots\ldots\ldots + \alpha_{ip}x_{ip} \quad (1)$$

where $\pi_i$ is the probability of the event of interest for the *jth* individual (Hosmer & Lemeshow, 2000). This model can be rewritten to express the probability itself:

$$\pi_i = \frac{exp(\alpha_{i0}+\alpha_{i1}x_{i1}+\alpha_{i2}x_{i2}+\alpha_{i3}x_{i3}\ldots+\alpha_{ip}x_{ip})}{1+exp(\alpha_{i0}+\alpha_{i1}x_{i1}+\alpha_{i2}x_{i2}+\alpha_{i3}x_{i3}\ldots+\alpha_{i1}x_{i1})} \quad (2)$$

The left-hand side, the logit, represents the logarithm of the odds that an event occurs, where odds are defined as the ratio of the probability of the event to that of its non-occurrence. This ratio simplifies the interpretation of the model coefficients as the change in odds associated with a one-unit change in predictors (Menard, 2002).

While binary logistic regression is well-suited for dichotomous outcomes, extending it to ordinal data introduces complexity due to the inherent order of categories. McCullagh (1980) highlighted that models for ordinal data leverage the ordinal nature by describing stochastic ordering, often requiring scores or assumptions of cardinality. Unlike multinomial logistic regression, which ignores order, ordinal models preserve the ranking, making them more appropriate for ordered categorical variables. Continuous development of these models remains crucial for accurate analysis of ordinal data in educational and social science research.

**Ordinal Logistic Regression (OLR)**

Ordinal Logistic Regression (OLR) is a statistical technique used when the dependent variable is ordinal, meaning its categories have a natural order but the intervals between categories are not necessarily equal (McCullagh, 1980). The independent variables in OLR can be categorical, interval, or ratio-scaled, allowing flexibility in modeling diverse types of predictors. When response categories possess an inherent order—such as levels of satisfaction, academic grades, or severity of an ailment—it is essential to incorporate this ordering into the model to fully utilize the available information (Agresti, 2010).

In OLR, the ordering is integrated directly into the specification of the logits, which represent the cumulative probabilities of the response variable. Specifically, the model estimates the log-odds of being at or below a particular category versus above it. This approach ensures that the ordinal nature of the data guides the modeling process, making the estimates more meaningful and interpretable (McCullagh, 1980).

In some cases, responses are continuous variables that are challenging to measure directly, prompting researchers to discretize the range into J ordinal categories, each associated with a probability ($\pi_j$). The discretization allows for modeling the probability of an observation falling within each category, facilitating analysis where direct measurement is impractical (Das & Rahman, 2011; Ombur & Gichuhi, 2011).

Two widely used models for ordinal data include the proportional odds model and the continuation ratio model. The proportional odds model assumes that the effect of

predictors is consistent across all thresholds of the response variable, simplifying interpretation and estimation (McCullagh, 1980). These models are particularly useful in fields like social sciences, medicine, and education, where ordered response data are common. divided into J ordinal categories with associated probabilities $\boldsymbol{\pi_1, \pi_2, \pi_3, \pi_4, \ldots \ldots . \pi_J}$. There are two commonly used models for this situation. (Das and Rahman, 2011, Ombur and Gichuhi, 2011)

**Cumulative Logit Model:** ,

The cumulative odds fr the jth response category is given by

$\frac{\boldsymbol{\pi_1+\pi_2+\pi_3+\pi_4,\ldots\ldots+\pi_J}}{\boldsymbol{\pi_{j+1}+\pi_{j+2}+\pi_{j+3}+\pi_{j+4}+\cdots\ldots+\pi_J}}$ where J is the total number of response categories.

The cumulative logit model is:

$$\log\left(\frac{\boldsymbol{\pi_1+\pi_2+\pi_3+\pi_4,\ldots\ldots+\pi_J}}{\boldsymbol{\pi_{j+1}+\pi_{j+2}+\pi_{j+3}+\pi_{j+4}+\cdots\ldots+\pi_J}}\right) = \alpha_{0j} + \alpha_{1j}x_{1j} + \alpha_{2j}x_{2j} + \alpha_{3j}x_{3j} \ldots + \alpha_{pj}x_{pj} \quad (4)$$

The odds ratio for each predictor is taken to be constant across all possible collapsing of the outcome variable. When a testable assumption is met, the odds ratio in a cumulative logit model are interpreted as the odds of being "lower" or "higher" on the outcome variable, across the entire range of the outcome. The wider applicability and intuitive interpretation of the cumulative logit model are two reasons for its being considered the most popular model for ordinal logistic regression

**Proportional Odds Model (POM)**

The Proportional Odds Model (POM) is a widely used extension of binary logistic regression tailored for ordinal response variables, where the categories have a natural order (Agresti, 2010). The model estimates the cumulative probabilities of being at or below a certain category, expressed through a series of logistic regressions. A key feature of POM is the proportional odds assumption, which posits that the effects of predictor variables are consistent across all thresholds of the response variable, meaning the regression coefficients are the same regardless of the category (Brant, 1990). Mathematically, the model is represented by the cumulative logit:

$$\log\left(\frac{\sum_{k=1}^{j}\pi_k}{\sum_{k+1}^{J}\pi_k}\right) = \alpha_{0j} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \cdots + \alpha_p x_p$$

where ( $\pi_k$ ) is the probability of the response being in category k, and $\alpha_{0j}$ are intercepts that depend on the category (j) (Mc Cullagh, 1980). The assumption implies that the

effects of predictors ($\alpha_1, \alpha_2\ \alpha_3 \cdots\cdots \alpha_p$) are invariant across categories, simplifying interpretation. The suitability of the POM can be evaluated by testing the proportionality assumption for each predictor, comparing it to models where the effects vary across categories. When the assumption holds, the model provides efficient and interpretable estimates for ordinal data (Brant, 1990; Agresti, 2010). The appropriateness of the model can be tested separately for each variable by comparing it to a cumulative odds model with the relevant parameter not depending on J.

## METHODOLOGY

**Link Function**:

The link function transforms cumulative probabilities to enable model estimation in ordinal regression (OR) analysis. Common link functions include logit, complementary log-log, negative log-log, probit, and Cauchit, each used to develop specific models. Choosing the most suitable link function depends on heuristics, as there's no universal method for selection. This transformation helps relate the probabilities to predictors, facilitating accurate modeling of ordinal data. The choice impacts the model's performance and interpretation, making understanding these functions crucial for effective analysis.

**Different Link Functions and Forms in Ordinal Regressions**

| s/n | Function | Form | Typical Application |
|---|---|---|---|
| 1 | Logit | $\mathbf{log}\left(\frac{\boldsymbol{F_x(x_i)}}{\mathbf{1}-\boldsymbol{F_x(x_i)}}\right)$ | Evenly distributed categories |
| 2 | Complementary log-log | $\mathbf{log}[(-\mathbf{log}\,\mathbf{1} - \boldsymbol{F_x(x_i)})]$ | Higher categories are more probable |
| 3 | Nagtive log-log | $-\mathbf{log}\big(\boldsymbol{F_x(x_i)}\big)$ | Lower categories are mor probable |
| 4 | Probit | $\emptyset^{-1}\boldsymbol{F_x(x_i)}$ | Normal distributed latent variable |
| 5 | Cauchit | $\mathbf{tan}\big(\boldsymbol{\pi}((\boldsymbol{F_x(x_i)} - \mathbf{0.5}))\big)$ | Outcome with many extreme values |

Hahn and Soyer (2010) found clear evidence that model fit can be improved by the selection of the appropriate link even in small data sets.

**Choice of Baseline Category:**

In ordinal regression, the baseline category (or reference group) is crucial for interpretation. For the dependent variable, such as final grades, first-class students are set as the baseline since they represent the desirable outcome. For other factors, the baseline

category is selected randomly because ordinal regression allows interpretation of factors regardless of this choice. Coefficients are derived from comparisons between categories, with the baseline category's coefficient set to zero, enabling meaningful interpretation of how other categories differ from the baseline in influencing the outcome.

**Odds Ratios and Model Interpretation**

Often it is not easy to directly interpret the model parameters $\alpha_1, \alpha_2\, \alpha_3 \cdots\cdots \alpha_p$ ar. Odds ratios (OR) often provide a much easier interpretations OR can be converted to its complement by dividing the OR by 1, e.g. OR (Category1/ category2) $= \frac{1}{OR\ for\ category\ 1}$

Interpreting model parameters directly can be challenging, so odds ratios (OR) often provide clearer insights. An OR compares the likelihood of an event occurring versus not occurring. For example, if the odds of a student receiving a particular grade are considered, the OR can be transformed into its reciprocal by dividing 1 by the OR, simplifying interpretation. In ordinal logistic regression, where student grades are ranked as Pass (1), Second Class Lower (2), Second Class Upper (3), and Distinction (4), we model odds for each category. For instance, $\vartheta_1$ is the odds of getting grade 1 versus higher grades, $\vartheta_2$ compares grades 1 or 2 to higher grades, and $\vartheta_3$ compares grades 1, 2, or 3 to higher grades. The highest grade, Distinction, doesn't have an associated odds because the cumulative probability of scoring up to this grade equals 1, making the odds undefined. These odds ratios help interpret how factors influence the likelihood of students achieving different grade levels, providing a more intuitive understanding of the model's results.

$$\vartheta_1 = \frac{prob(grade\ 1)}{prob.(grade\ greater\ than\ 1)}$$
$$\vartheta_2 = \frac{prob(grade\ 1\ or\ 2)}{prob.(grade\ greater\ than\ 2)}$$
$$\vartheta_3 = \frac{prob(grade\ 1{,}2\ or 3)}{prob.(grade\ greater\ than\ 3)}$$
$$\vartheta_4 = \frac{prob(grade\ 1{,}2{,}3{,}or\ 4)}{prob.(grade\ greater\ than\ 4)}$$

And the research is limited to the fourth model

$$\boldsymbol{\vartheta_4 = \frac{prob(grade\ 1, 2, 3, or\ 4)}{prob.\ (grade\ greater\ than\ 4)}}$$

**Test of Parallelism:**

The test of parallelism in ordinal regression assesses whether the regression coefficients are consistent across all categories. A fundamental assumption here is that the proportional odds, or parallelism, holds true—meaning the relationship between predictors and the outcome remains constant across categories (Agresti, 1992). When this assumption is

violated, it is advisable to switch to multinomial regression, which estimates separate coefficients for each category, offering more flexibility (Williams, 2012). If the null hypothesis of parallelism is rejected at a specified significance level, it suggests that either the chosen link function may be inappropriate or that the relationships between independent variables and logits differ across categories (O'Connell, 2006). Nonetheless, the test for proportional odds is often considered anti-conservative, frequently leading to rejection of the assumption, especially with numerous predictors or large samples (Brant, 1990; Allison, 1999). This tendency underscores the importance of carefully evaluating the assumption before proceeding with ordinal regression models (Cohen, 1993; Agresti, 2013).

**Parameter Estimates of the Coefficients:**

Parameter estimates provide detailed insight into how explanatory variables relate to the outcome in ordinal regression models (Mc Cullagh, 1980; Agresti, 2013). Typically, threshold coefficients are not interpreted individually; instead, they function as intercepts indicating the logit point where individuals are predicted to shift into higher categories (O'Connell, 2006). While direct interpretation of these thresholds is often unnecessary, understanding them helps clarify how the model estimates probabilities across categories and can assist in assessing model fit, including measures like $R^2$ (Nagelkerke, 1991). These coefficients essentially mark the boundaries between categories, providing context for the effects of predictors on the likelihood of falling into each outcome level (Lemeshow et al., 1993). Recognizing their role enhances comprehension of the overall model dynamics (Cohen, 1993).

**Measuring Strength of Association**

Assessing the strength of association between a dependent variable and predictor variables in logistic regression differs from linear regression. While $R^2$ measures are common in linear models, their interpretation in logistic regression is less straightforward (Hosmer & Lemeshow, 2000). Several pseudo-$R^2$- statistics provide alternative measures of fit (Menard, 2001). (Cox & Snell, 1989), (Nagelkerke, 1991), and (McFadden, 1974) are frequently used. These statistics, calculated using log-likelihood functions, estimate the proportion of variance explained by the model. However, their interpretation requires careful consideration, as they are not directly comparable in linear regression. Differences in their ranges and calculation methods (such as the denominator in Nagelkerke's ) further complicate their interpretation.. The three commonly used statistics are:

1. Cox and Snell $R^2$ : $R^2_{CS} = 1 - \left(\frac{L(B^{(0)})}{L(\hat{B})}\right)^{\frac{2}{n}}$

2. Nagelkerke's $R^2$ : $R_N^2 = \frac{R_{CS}^2}{1-L(B^{(0)})^{\frac{2}{n}}}$

3. McFadden's $R^2$ : $R_M^2 = 1 - \left(\frac{L(\hat{B})}{L(B^{(0)})}\right)$

Where

$L(\hat{B})$ is the log-likelihood function for the model with estimated parameters , $L(B^{(0)})$ is the log-likelihood with just the thresholds, and $\boldsymbol{n}$ is the number of cases (sum of all weights).

**DESCRIPTIVE ANALYSIS**

Before constructing the model, a descriptive analysis was conducted to explore the data, as shown in the appendix. Figures 1 to 4 present cumulative percentage plots of four factors:—department, mode of entry, sex, and age—against final grades. Figure 1 indicates that students in the Business Administration and Management (BAM) department have the highest likelihood of graduating with distinction. Figure 2 reveals that students admitted with a good or excellent academic background are more likely to attain higher final grades than those with poorer academic backgrounds. Regarding sex, the points for males generally fall below those for females, but there are overlaps, suggesting sex may not significantly influence the odds of achieving a final grade (Field, 2013). Figure 4 shows that students older than 20 at entry tend to have higher cumulative percentages of distinction, implying that younger students (≤20 years) have lower odds of achieving top grades. This observation aligns with existing research indicating age can influence academic performance (Tinto, 1993; Yorke & Longden, 2024). Overall, these figures provide valuable initial insights, highlighting potential relationships between demographic factors and final academic outcomes, which warrant further statistical testing in the model-building process.

**Choosing Appropriate Link Function**

Choosing an appropriate link function is a critical step in modeling ordinal categorical variables, such as students' final grades. The selection process relies heavily on understanding the relationship between the dependent variable and the predictors, as well as examining the data's underlying structure. In this context, the plot of the final grade against the frequency of occurrence provides valuable insights, guiding the choice of the most suitable link function. The figure below indicates that the PROBIT link function is appropriate for modeling the ordinal nature of final grades in this dataset. It suggests that the relationship between the predictors and the outcome aligns well with the assumptions

underlying the PROBIT model, which is often preferred when the latent variable underlying the ordinal response is assumed to follow a normal distribution.

When evaluating the performance of different models, statistical significance and goodness-of-fit measures are essential. According to the results from Tables 2 and 3, the model predicting final grade based on sex does not significantly outperform the baseline intercept-only model. This implies that including sex as a predictor does not enhance the model's ability to predict final grades beyond what is achieved by simply using the marginal probabilities of the outcome categories. In other words, sex alone does not provide meaningful information for predicting final grades in this dataset. Conversely, all other models—those incorporating variables such as age at entry, department, and mode of entry—show a significant improvement over the baseline. This indicates that these factors contribute valuable information, enabling the models to predict final grades more accurately than a naive approach based solely on marginal probabilities.

Despite the low R-squared ($R^2$) values reported in Table 4, these figures are the best available under the PROBIT link function for this particular dataset. It is important to recognize that $R^2$ values in the context of ordinal regression and models employing non-linear link functions tend to be lower and less directly interpretable than in linear regression. Different application domains have varying standards for what constitutes a **"good" $R^2$;** thus, these metrics should be viewed as comparative rather than absolute measures. When comparing models fitted to the same data, the one with the highest **$R^2$** is generally considered superior. Nonetheless, the pseudo $R^2$ values provide a useful, albeit approximate, gauge of relative model performance. It is worth noting that in regression analysis, the **$R^2$** statistic remains more straightforward and interpretable than pseudo $R^2$ measures, which often require more cautious interpretation.

Further analysis of the model fit is provided by Table 5, which tests the proportional odds assumption—a key consideration in ordinal regression modeling. The results show that the model predicting final grade based on sex violates this assumption, indicating that the relationship between predictors and the response varies across the thresholds of the ordinal outcome. This violation suggests that a single set of coefficients may not adequately capture the effect of sex across all grade categories. In contrast, the model incorporating age at entry, department, and mode of entry does not reject the proportional odds assumption, implying that these variables have a consistent effect across different thresholds of final grades. This consistency enhances the interpretability and stability of the model for these predictors.

The coefficients and odds ratios (OR) associated with the predictors further elucidate the relationships. For instance, Table 6 reports the coefficient for male students as 0.0331. Exponentiating this coefficient yields an OR of approximately 1.0315, with females serving as the baseline category. Reversing the sign of the coefficient provides the OR for males relative to females, which is approximately 0.9695. The **P – value** associated with

the sex coefficient (0.728) indicates that sex is not a statistically significant predictor of final grades, reinforcing the earlier conclusion that gender does not substantially influence academic attainment in this model. The odds ratio close to 1 suggests similar odds for males and females achieving higher final grades, with neither group showing a significant advantage.

In addition, Table 7 shows that age at entry influences final outcomes; students aged 20 or below at entry have approximately 1.35 times higher odds of graduating with distinction compared to those older than 20. This odds ratio of 1.3499 indicates that younger students are somewhat more likely to attain higher grades, though the significance of this effect depends on the associated P-value and confidence intervals. Such findings highlight the importance of considering demographic factors when modeling academic achievement, but also emphasize the need for appropriate statistical tests to confirm these relationships.

In conclusion, selecting the appropriate link function is vital for accurate and interpretable ordinal regression modeling. The graphical evidence supports the use of the PROBIT link function, which aligns well with the data's distributional assumptions. While some predictors like sex do not significantly improve the model, others such as mode of entry and age at entry contribute meaningful information. Despite the low $R^2$ values, these models offer valuable insights into the factors influencing students' final grades, emphasizing the importance of choosing the correct link function and verifying model assumptions to ensure robust and meaningful results.

## CONCLUSIONS

From the analyses conducted, several key conclusions can be drawn regarding the factors influencing students' final grades.

1. Among all the variables examined, only sex appears to have no dependency on the final grade; it does not significantly influence academic outcomes.
2. Models incorporating variables such as age at entry, department, and mode of entry demonstrate superior predictive ability compared to simply relying on the marginal probabilities of outcome categories, indicating their meaningful contribution to understanding final grades.
3. The analysis reveals that sex contributes the least to explaining the variation in final grades, whereas the grade of entry—such as academic background—provides the most substantial explanation, emphasizing its importance in predicting academic success.
4. The proportional odds assumption is only rejected for the sex variable, implying that the relationship between sex and final grade varies across different grade thresholds. For all other factors, this assumption holds, supporting the consistency of their effects across categories.

5. Notably, sex is also found to be statistically insignificant in predicting final grades, reinforcing the conclusion that gender does not play a decisive role in academic achievement at graduation.
6. Consequently, the data suggest that neither male nor female students have a consistent advantage in attaining higher final grades.
7. Additional findings indicate that students aged 20 or below at entry are approximately 1.35 times more likely to graduate with distinction compared to older students.
8. Moreover, students with academic advantages are over three times more likely to graduate with distinction than those with lesser academic preparation.
9. Similarly, students with a strong academic background are more than one and a half times more likely to graduate with distinctions or a second-class upper division compared to students with poorer academic records.
10. Overall, these results highlight that demographic factors like sex are not determinants of academic success, whereas factors such as age at entry and academic background significantly influence final outcomes.

Appendix

**Table 1: Data Summary**

| S/N | | VARIABLE | CODE | N | Marginal Percentage |
|---|---|---|---|---|---|
| 1 | Final Grade | Pass | [CGPA :2.00 – 2.49] | 67 | 21.4% |
| | | Lower credit | [CGPA :2.50 – 2.99] | 345 | 49.4% |
| | | Upper credit | [CGPA: 3.00 – 3.49] | 305 | 53.2% |
| | | Distinction | [CGPA: 3.50 – 4.00] | 45 | 2.8% |
| 2 | Department | Accounting | [dept = 1] | 234 | 12.3% |
| | | Agriculture | [dept = 2] | 169 | 10.2% |
| | | ABEET | [dept = 3] | 125 | 9.6% |
| | | Building Tech | [dept = 4] | 40 | 18.6% |
| | | BAM | [dept = 5] | 450 | 28.9% |
| | | Civil Engineering | [dept = 6] | 63 | 4.2% |
| | | Computer Engine | [dept = 7] | 134 | 9.8% |
| | | Computer Science | [dept = 8] | 345 | 27.8% |
| | | EEE | [dept = 9] | 168 | 10.5% |
| | | PAD | [dept = 10] | 568 | 36.3% |
| | | Statistics | [dept = 11] | 145 | 10.1% |
| | | SLT | [dept = 12] | 784 | 54.7% |
| | | OTM | [dept = 13] | 702 | 52.5% |
| 3 | Age at entry | $\leq 20\ years$ | [Age = 1] | 1650 | 67.4% |
| | | $\geq 20\ years$ | [Age = 2] | 1245 | 32.6% |
| 4 | Sex | Male | [sex = 1] | 1870 | 56.8% |
| | | Female | [sex = 2] | 1243 | 43.2% |
| 5 | Mode of performance | Good Academic background | [mode =1] | 1745 | 43.2% |
| | | Poor Academic background | [mode = 2] | 1945 | 56.8% |

## Table 2: chi-square Test of Independence

| s/n | Factor | Pearson $x^2$ value | $df$ | p-value | Decision |
|---|---|---|---|---|---|
| **1** | **Sex*final grade** | 10.71 | 4 | 0.030 | Sex does not depend on the final grade obtained at 2.7% significance level. |
| **2** | **Department* final grade** | 129.80 | 45 | .001 | Department is dependent on the final grade |
| **3** | **Age at entry*final grade** | 24.48 | 4 | .001 | Age at entry is dependent on the final grade. |
| **4** | **Mode of performance * final grade** | 108.04 | 8 | .001 | Mode of performance is dependent on the final grade |

## Table 3: ORDINAL REGRESSION
## Model fitting Information

| s/n | Factor | Model | -2Log likelihood | Chi-square | $df$ | $sig.$ |
|---|---|---|---|---|---|---|
| **1** | **Sex*final grade** | Intercept only | 47.90 | 0,132 | 1 | 0.823 |
| | | Final | 46.99 | | | |
| **2** | **Department* final grade** | Intercept only | 231.80 | 57.06 | 10 | 0.000 |
| | | Final | 178.01 | | | |
| **3** | **Age at entry*final grade** | Intercept only | 60.54 | 12.10 | 1 | 0.001 |
| | | Final | 49.56 | | | |
| **4** | **Mode of performance * final grade** | Intercept only | 163.12 | 78.01 | 2 | 0.000 |
| | | Final | 63.98 | | | |

## Table 4: measure of Determination

| s/n | model | Nagelkerke $R^2$ | Cox and Snell$R^2$ | Mc Fadden $R^2$ |
|---|---|---|---|---|
| **1** | **Sex*final grade** | **0.001** | **0.000** | **0.000** |
| **2** | **Department* final grade** | **0.085** | **0.089** | **0.040** |
| **3** | **Age at entry*final grade** | **0.017** | **0.019** | **0.006** |
| **4** | **Mode of performance * final grade** | **0.131** | **0.119** | **0.074** |

## Table 5: Test of Parallel Lines

| s/n | Factor | Model | -2Log likelihood | Chi-square | $df$ | $sig.$ |
|---|---|---|---|---|---|---|
| **1** | **Sex*final grade** | Null Hypothesis | 47.801 | 11,132 | 3 | 0.003 |
| | | General | 56.996 | | | |
| **2** | **Department* final grade** | Null Hypothesis | 181.805 | 7.069 | 23 | 0.239 |
| | | General | 168.016 | | | |
| **3** | **Age at entry*final grade** | Null Hypothesis | 48.540 | 2.102 | 4 | 0.151 |
| | | General | 45.564 | | | |
| **4** | **Mode of performance * final grade** | Null Hypothesis | 63.121 | 1.901 | 6 | 0.585 |
| | | General | 63.980 | | | |

## Table 6: Parameter Estimates (Sex*Final Grade)

| | | Estimate | Std. Error | Wald | $df$ | $sig.$ | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Lower** | **Upper** |
| Threshold | [CGPA :2.00 – 2.49] | **-2.165** | **0.139** | **240.405** | **1** | **0.000** | **-2.540** | **-1.781** |
| | [CGPA :2.50 – 2.99] | **-0.684** | **0.0689** | **94.630** | **1** | **0.000** | **-0.887** | **-0.599** |
| | [CGPA: 3.00 – 3.49] | **0.423** | **0.727** | **31.211** | **1** | **0.000** | **0.268** | **0.468** |
| | [CGPA: 3.50 – 4.00] | **2.560** | **0.123** | **271.654** | **1** | **0.000** | **1.857** | **2.431** |
| Location | **[sex = 1]** | **0.045** | **0.087** | **0.131** | **1** | **0.695** | **0.212** | **0.215** |
| | **[sex =2]** | **0** | **-** | **-** | **0** | **-** | **-** | **-** |

**Odds Ratio**

| s/n | Gender | |
|---|---|---|
| 1 | **Female – Male** | **1.0452** |
| 2 | **Male – Female** | **0.9899** |

**Table 7: Parameter Estimates (Age at entry*Final Grade)**

| | | Estimate | Std. Error | Wald | *df* | *sig.* | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Lower** | **Upper** |
| Threshold | [CGPA :2.00 – 2.49] | **-2.065** | **0.145** | **206.105** | **1** | **0.000** | **-2.340** | **-1.741** |
| | [CGPA :2.50 – 2.99] | **-0.484** | **0.089** | **62.430** | **1** | **0.000** | **-0.687** | **-0.499** |
| | [CGPA: 3.00 – 3.49] | **0.553** | **0.089** | **57.081** | **1** | **0.000** | **0.488** | **0.768** |
| | [CGPA: 3.50 – 4.00] | **2.236** | **0.123** | **261.154** | **1** | **0.000** | **2.037** | **2.634** |
| Location | **[Age = 1]** | **0.285** | **0.097** | **12.131** | **1** | **0.001** | **0.112** | **0.515** |
| | **[Age =2]** | **0** | **-** | **-** | **0** | **-** | **-** | **-** |

**Odds Ratio**

| s/n | Gender | |
|---|---|---|
| 1 | **> 20 years /≤ 20 years** | **1.3452** |
| 2 | **≤ 20 years />20 years** | **0.7094** |

**Table 8: Parameter Estimates (Mode of Performance*Final Grade)**

| | | Estimate | Std. Error | Wald | *df* | *sig.* | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Lower** | **Upper** |
| Threshold | [CGPA :2.00 – 2.49] | **-1.795** | **0.195** | **216.285** | **1** | **0.000** | **-3.360** | **-2.771** |
| | [CGPA :2.50 – 2.99] | **-1.884** | **0.149** | **82.400** | **1** | **0.000** | **-1.687** | **-1.119** |
| | [CGPA: 3.00 – 3.49] | **-1.153** | **0.139** | **7.041** | **1** | **0.300** | **-0.438** | **0.168** |
| | [CGPA: 3.50 – 4.00] | **1.636** | **0.173** | **95.154** | **1** | **0.000** | **1.337** | **1.834** |
| Location | **[mode = 1]** | **-1.225** | **0.197** | **52.181** | **1** | **0.001** | **-1.412** | **-0.715** |
| | **[mode =2]** | **-0.389** | **0.143** | **8.315** | **0** | **0.004** | **-0.713** | **-0.131** |

**Odds Ratio**

| s/n | performance | |
|---|---|---|
| **1** | Good academic background | 0.3452 |
| **2** | Poor academic background | 0.6094 |

## References.


Academic Policy (2019/2020) Federal Polytechnic, Bali Pp 44 – 45

Agresti, A. (1997). "Modeling ordinal response data". Journal of the Royal Statistical Society: Series B (Methodological), Vol.59 No.(3), Pp 576 – 582.

Agresti, A. (2010). "Analysis of ordinal categorical data". Wiley Series in Probability and Statistics.

Brant, R. (1990). "Logistic regression modeling of ordinal response variables". Journal of the Royal Statistical Society: Series B (Methodological), Vol.52 No.(4), Pp 229 – 242.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. "Biometrics", Vol. 46 Vol.(4), Pp1171 – 1178.

Cohen, J. (1993). "A power primer". Psychological Bulletin, Vol.112 No.(1), Pp 155 – 159.

Cox, D. R., & Snell, J. (1989). "Analysis of binary and polytomous response data". Chapman and Hall.

Das, S., & Rahman, M. (2011). "Ordinal data modeling: An overview". Journal of Statistical Computation and Simulation, Vol.81 No.(4), Pp 483 – 502.

Das, S., & Rahman, M. (2011). "Ordinal Logistic Regression Model". Journal of Statistical Planning and Inference, Vol.141 no.(3), Pp 1214 – 1227.

Ebenuwa-Okoh, B. (2010). "Factors influencing academic performance among Nigerian undergraduates". Journal of Education and Social Research, Vol.1 No.(4), Pp 22 – 30.

Egbule, C. L. (2004). "Correlates of academic achievement among Nigerian polytechnic students". Nigerian Journal of Educational Research and Evaluation, Vol.8 No.(2), Pp 25 – 34.

Exams & Record (2022-2024), Academic Registry Unit, Registry, Federal Polytechnic, Bali.

Field, A. (2013). "Discovering Statistics Using IBM SPSS Statistics" (4th ed.). Sage Publications.

Hahn, J., & Soyer, R. (2010). "Choosing link functions in ordinal regression". Communications in Statistics - Simulation and Computation, Vol.39 No.(10), PP1827 – 1839.

Hosmer, D. W., & Lemeshow, S. (2000). "Applied Logistic Regression" (2nd ed.).Wiley.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). "Applied Logistic Regression". (3rd ed.). Wiley.

Lemeshow, S., Hosmer, D. W., Klar, J., & Lwanga, S. K. (1993). "Adequacy of sample size in health studies". Wiley.

Long, J. S., & Freese, J. (2018). "Regression models for categorical dependent variables using Stata". (3rd ed.). Stata Press.

McCullagh, P. (1980). "Regression models for ordinal data". Journal of the Royal Statistical Society: Series B (Methodological), Vol. 42 No.(2), Pp109 – 127.

McFadden, D. (1974). "Conditional logit analysis of qualitative choice behaviour". In P. Zarembka (Ed.), Frontiers in Econometrics (Pp 105 – 142). Academic Press.

Menard, S. (2002). "Applied logistic regression analysis" (2nd ed.). Sage Publications.

Nagelkerke, N. J. D. (1991). "A note on a general definition of the coefficient of determination". Biometrika, Vol. 78 No.(3), Pp 691 – 692.

O'Connell, A. (2006). "Ordinal logistic regression". In Applied ordinal logistic regression. New York: Springer.

O'Connell, A. A. (2006). "Logistic regression models for ordinal response variables". Journal of Modern Applied Statistical Methods, Vol.5 No.(3), Pp 390 – 404.

O'Connell, A. A. (2021). "Logistic regression for ordinal response variables". Advances in Social Science Methodology, Vol.17 No.(3), Pp 45 – 62.

Ombur, R. K., & Gichuhi, J. W. (2011). Application of ordinal logistic regression in modeling ordered categorical data. "African Journal of Statistics", Vol.3 no(2), Pp 123 – 134.

Ombur, W. K., & Gichuhi, P. N. (2011). "Application of ordinal logistic regression in social sciences". International Journal of Social Science Studies, Vol.3 No.(2), Pp 112 – 122.

Petersen, K., & Svanes, T. (1994). "Ordinal logistic regression: An introduction". Scandinavian Journal of Statistics, Vol.21 No.(2), Pp133 – 148.

Tinto, V. (1993). "Leaving college: Rethinking the causes and cures of student attrition". University of Chicago Press.

Tinto, V. (1997). "Classics in student retention". Journal of College Student Retention, Vol.1 No.(1), Pp7 – 22.

Ugoji, E. O. (2008). “Factors affecting academic performance of students in Nigerian polytechnics”. Journal of Education and Practice, Vol3 No.(12), Pp 45 – 52.

Ukpong, D. I. (2007). “Determinants of academic performance among Nigerian undergraduates”. African Journal of Educational Studies, Vol. 4 No. (1), Pp 78 – 85.

Williams, R. (2012). “Understanding and interpreting the results of the proportional odds model”. The Stata Journal, Vol.12 No.(2), Pp 211 – 229.

Williams, R. (2019). “Ordinal logistic regression: Model interpretation”. Journal of Statistical Software, Vol.88 NO.(5), Pp 1 – 20.

Yorke, M., & Longden, B. (2024). “Student success and retention in higher education”. Journal of Higher Education Policy and Management, Vol.46 No.(2), Pp 123 – 135.