

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(12): 143-153
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 26-10-2025
Accepted: 29-11-2025

JA Otunmeruke
Southern Delta University,
Ozoro, Delta State, Nigeria

M Osolo
Chukwuemeka Odumegwu
Ojukwu University, Uli,
Anambra State, Nigeria

A Okogi
Southern Delta University,
Ozoro, Delta State, Nigeria

N Diamond
Southern Delta University,
Ozoro, Delta State, Nigeria

Corresponding Author:
JA Otunmeruke
Southern Delta University,
Ozoro, Delta State, Nigeria

Modelling the social and economic effects of road traffic accidents in Delta State, Nigeria: A multivariate analysis approach

JA Otunmeruke, M Osolo, A Okogi and N Diamond

DOI: <https://www.doi.org/10.22271/math.2025.v10.i12b.2218>

Abstract

This study presents a statistical modeling approach to examining the socio-economic effects of road traffic accidents in Delta State, Nigeria. Using quarterly secondary data from 2015 to 2024 from the Federal Road Safety Corps and the National Bureau of Statistics for the period 2015-2024. The study employs multiple regression analysis to examine how explanatory variables predict the number of cases, relationship between accident-related variables (cases, people involved, casualties) and socio-economic impacts. The findings indicate that all the explanatory variables such as accident counts, injured, killed and people involved are statistically significant with $p < 0.05$. The 'effect' of each variable on the mean response when other variables are held constant shows the existence of a linear relationship. The model recorded R^2 (0.943), this explains 94.3% of the variance in accident cases; indicating the combination of injured, killed and people involved variable strongly predicts accident case numbers. The adjusted R^2 (0.939) confirms a very good model fit. The 'effect' of each variable on the mean response when other variables were held constant shows the existence of a relative linear relationship. The results shows that Principal Component one alone captures almost all the information (91.9%) in the three original predictors. While PC2 adds only 6.1%, meaning together PC1 and PC2 represent 98% of the explained variance. This confirms that the predictors are strongly correlated, which principal component analysis has resolved by creating uncorrelated axes. Hence, the predictors of road accident consequences provide insight into health, policy intervention strategies, and road safety management.

Keywords: Multiple Regression, Statistical Diagnostics, Impact, Road Traffic Accidents, Principal Component Analysis

1. Introduction

1.1 Background to the Study

Road Traffic Accidents (RTAs) are among the leading causes of mortality and socioeconomic disruption globally. In Nigeria, the challenge is magnified by inadequate infrastructure and enforcement, especially Delta State, a hub of commercial and intercity transport, experiences frequent RTAs, resulting in deaths, injuries, and economic displacement. Yet, limited research has adopted formal statistical models to quantify how these incidents affect households' livelihoods and education. The Federal Road Safety Corps (FRSC) has consistently reported alarming statistics on road accidents in Nigeria, with Delta State featuring prominently due to its strategic location and economic activities (FRSC, 2024). These accidents not only result in immediate physical injuries and fatalities but also have far-reaching consequences for families, including economic hardship, psychological trauma, and social disruption.

A quarterly analysis of accident data during this period reveals seasonal variations, such as increase crashes during festive seasons and rainy months, which are critical for designing targeted interventions. The economic implications of these accidents are profound, as families often face medical expenses, loss of income, and caregiving burdens, exacerbating poverty in a region with limited social safety nets.

Delta State's road safety challenges are compounded by systemic issues, including poor road maintenance and inadequate emergency response systems. The FRSC's annual reports

highlighted that road infrastructure deficiencies, such as potholes and lack of signage, contributed significantly to accidents in the state (FRSC, 2018). These environmental factors, combined with vehicular issues like poorly maintained trucks and motorcycles, create a complex web of causation that requires localized research. By focusing on Delta State, this study addresses a gap in the literature, as most existing studies adopt a national perspective, often overlooking state-specific dynamics (Oluwasanmi, 1993).

The socio-economic context of Delta State further amplifies the impacts of road accidents. As a region reliant on oil and trade, the state sees heavy traffic from commercial vehicles, which are frequently involved in crashes due to overloading and mechanical failures, as noted in FRSC records (FRSC, 2024). Families affected by these accidents often face immediate financial crises, particularly when victims are primary earners. The ripple effects include disrupted education for children, increased household debt, and strained community networks. This study focuses on quarterly data from 2015 to 2024 that uncover peak accident periods and their correlation with economic activities, providing a foundation for targeted road safety campaigns and support programs for affected families (Atubi, 2012)^[7]. Ogunbodede, (2008)^[13] examined RTAs in Nigeria and identified poor road conditions and human errors as key causes. Oni *et al.* (2011)^[18] analyzed accident trends in Lagos using time series and reported a significant correlation between traffic volume and accidents. Okafor and Eke (2017)^[14] used regression analysis to model accident fatality rates in Enugu and found income and road type to be significant predictors.

Aluko *et al.* (2019)^[6] evaluated the impact of RTAs on productivity loss in Southwest Nigeria, showing a strong relationship between accident involvement and loss of income. FRSC (2021) reported that over 5,000 accidents occurred nationwide in 2020 alone, with economic costs estimated in billions. WHO (2022) stated that road injuries are among the top 10 causes of death in Nigeria, highlighting the need for data-driven policy. Uzundu and Nduka (2020)^[19] analyzed secondary data on RTAs in the South-South region and suggested stronger enforcement of traffic laws and improvement in emergency response.

Okonkwo and Adebajo (2019)^[15] mapped crash locations but did not explore predictive models. Adebayo and Nwachukwu (2020)^[2] linked enforcement activities to accident reduction via simple correlation. Oluwasanmi *et al.* (2021)^[16] used time series analysis to track fluctuations. However, few studies combined socioeconomic variables with diagnostics and rigorous regression frameworks. However, statistical modeling approaches to transport safety have been covered extensively in international literature (e.g., Negative Binomial models, Poisson regression for crash counts, and logistic regression for binary outcomes) but not localized. This study addresses this gap by applying multivariate analysis to examine how variables such as: “Injured, Killed and People Involved or Casualties” predict the number of cases that influence on socioeconomic outcomes.

Statement of the Problem

Despite increasing efforts by road safety authorities, road traffic accidents remain prevalent in Delta State, causing loss of life, economic disruption, and social dislocation. This persistent high incidence of road accidents in State poses a significant threat to public safety and family welfare. Nevertheless, limited empirical research has quantified the

direct relationship between accident rates and outcomes like school dropout using statistical modeling. Most existing studies are descriptive and do not employ predictive models to understand these dynamics. This study addresses this gap by applying multiple regression, logistic analysis, and PCA to determine the extent to which road accidents influence social and economic outcomes using secondary data.

1.3 Aim and Objectives of the Study

The aim of this study is to model the social and economic effects of road traffic accidents in Delta State using multivariate analysis approach. Specific objectives are to:

1. Examine the trend of road traffic accidents from 2015 to 2024.
2. Examine how variables such as: “Injured, Killed and People Involved or Casualties” predict the number of cases.
3. Identify the effect of accidents on school dropout and job loss.
4. Evaluate the relationship between road accidents and family income levels.
5. Recommend policy interventions based on statistical findings.

1.4 Significance of the Study

This study provides evidence-based insights into the social and economic impacts of road traffic accidents using statistical modeling. It contributes to the literature on transport safety and social policy, and guides policymakers, urban planners, and healthcare agencies in designing interventions which would alleviate the negative effects on the individual households, communities, State and Nigeria at large.

1.5 Scope of the Study

The study is limited to Delta State, Nigeria, and uses secondary data from 2015 to 2024. It focused on quarterly variables such as accident counts, casualties, income levels, school dropout, and job loss.

2. Methodology

2.1 Research Design

This study adopts a quantitative research design, using secondary data to model the effects of road traffic accidents through multiple regression, logistics analysis and principal component analysis. Delta State was randomly selected among the six states (Akwa-Ibom, Bayelsa, Cross River, Delta, Edo, and Rivers) that comprises of South-South Geographical region using simple random sample method.

2.2 Sources of Data

Secondary data were sourced from the Federal Road Safety Corps (FRSC), National Bureau of Statistics (NBS), and other reliable publications. The data covers the period 2015-2024.

2.3 Variables and Measurement

The study measured accident counts, people involved, injured, killed, school dropout, job loss and family income.

- **Dependent Variables:** School dropout, Job loss, Family income
- **Independent Variables:** Accident count, People involved, Casualties (injured, killed)

2.4 Population and Sampling

Since secondary data is used, the population consists of recorded accident cases and associated social and economic indicators for Delta State. The full dataset for the stated period was analyzed, thus eliminating the need for sampling.

2.5 Method of Data Analysis

Descriptive statistics (mean, trend graphs) were used to explore the data. Multiple regression analysis was employed to assess the impact of independent variables on social and economic outcomes. SPSS version 26 software was adopted for computation.

2.6 Multiple Regression Assumptions

The assumptions that guided this analysis include:

1. **Linearity:** The relationship between predictors and the outcome variables is linear.
2. **Independence of Errors:** Residuals are not correlated across observations. Durbin-Watson statistic (ideal range 1.5 - 2.5).
3. **Homoscedasticity:** The variance of residuals remains constant across levels of independent variables.
4. **Normality of Residuals**
5. **Multicollinearity:** No multicollinearity means independent variables are not highly correlated with one another. VIF < 5 considered acceptable.

6. **No Measurement Errors:** All variables are accurately measured and recorded.

7. **Correct Model Specification:** All relevant predictors are included and irrelevant ones excluded.

These assumptions were assessed using diagnostic tools such as residual plots, variance inflation factor (VIF), and normality tests within SPSS.

2.7 Ethical Considerations

As secondary data is used, there are minimal ethical risks. However, all data were handled with confidentiality and used solely for this study with intention of evidence-based policy formulation.

2.8 Data Analysis Plan

The data analysis involved cleaning, handling of missing values, outliers, and ensured consistency. The modelling of "Road Traffic Accidents" looks at the linear, non-linear, and possible logistic regression analysis to test the accuracy, robustness and predictability of the methodology's appropriateness.

Model: In Multiple Linear Regression, the relationship between the dependent variable and independent variables is expressed as a linear combination of the variables. The model is written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Where:

Y = Outcomes

β_0 = Intercept (constant)

$\beta_1 + \beta_2 + \beta_3 + \dots + \beta_n$ = Coefficient for each variable

$X_1, X_2, \dots, X_n$ = Values of each independent variable

e = Error term

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon$$

$$\text{PredictedCases} = \beta_0 + \beta_1(\text{AccidentCount}) + \beta_2(\text{Injured}) + \beta_3(\text{Killed}) + \beta_4(\text{People_Involved}) + \epsilon_i$$

The regression coefficients: The coefficient for any one of the variables under study estimates how much the mean of the response variable change by one unit, when all other variables are held constant.

Model Diagnostics and Validation: The diagnostic and validation test were conducted

(e.g., residual analysis for linear and non-linear regression, goodness-of-fit for logistic regression).

2.8.2 Goodness of Fit Measures:

- R^2 : This measure indicates the proportion of variance in the dependent variable explained by the model. A higher (R^2) suggests a better fit.
- Adjusted R-squared: This adjusted (R^2) for the number of predictors in the model, penalized for unnecessary complexity.
- In interpreting our model, we focused on R^2 , Coefficients (β), and significance (p-values). The shape of the curve using interaction or squared terms aided the interpretation.

2.8 Limitations of the Study

The study relies on the accuracy and completeness of secondary data. Lack of access to disaggregated or household-level data may limit the depth of analysis.

3.1 Data Analysis and Results

This chapter presents the analysis of data obtained from secondary sources. It includes descriptive statistics, trend line graph, and regression results that test the relationship between road traffic accident variables and socioeconomic outcomes. Principal component analysis (PCA) was used to eliminate or reduce multicollinearity in the data.

3.2 Descriptive Statistics

The mean of the road traffic accident counts show that killed has 38.26 and cases records 58.53 with their standard

deviations to be 18.52 and 35.92 respectively.

a3.3 Trend Line Graph

The trend line graph illustrates the quarterly/yearly unstable behaviour and changes in accident numbers, casualties, and identified some visible patterns or shifts. We observed that 2019 witnessed the peak period of RTAs under the period of our investigation. This trend could have been due to increase in travelling rates due to fear of the COVID-19 pandemic in RTAs in Delta State. The graph revealed a statistically significant upward trend in RTAs during festive periods, weekends, and wet seasons. The leading causes most likely could be speeding, poor road infrastructure, mechanical failure, and driver distraction (uses of phones, alcohol). The line graph is linear.

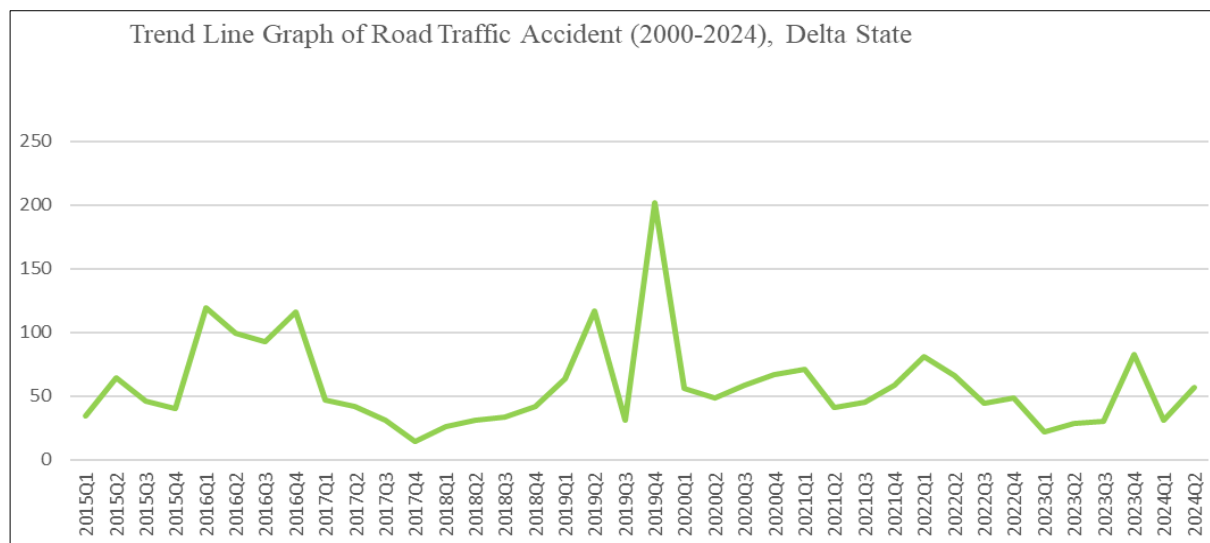


Fig 1: Trend line graph of road traffic accidents

Explanatory Variables Predicting Cases

The researcher also examined how variables such as: “Injured, Killed and People Involved” predict the number of cases or casualties. In this model, we have the following:

- **Dependent Variable:** Cases
- **Independent Variables:** Injured, Killed and People Involved

Regression Equation

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Y = Accident count

X₁ = Injured

X₂ = Killed

X₃ = People Involved

$$\text{PredictedCases} = \beta_0 + \beta_1(\text{Injured}) + \beta_2(\text{Killed}) + \beta_3(\text{People_Involved})$$

R Correlation between observed and predicted values of Cases

- **R Squared:** Proportion of variance in cases explained by the model (e.g. 0.94= 94%)
- **Adjusted R Square:** Corrected R² for the number of predictors; good for model comparison.

Table 1: Model Summary

R	R Squared	Adjusted R
0.971	0.943	0.939

R² (0.943): The model explains 94.3% of the variance in accident cases. This means the combination of injured, killed and people involved strongly predicts accident case numbers. Adjusted R² (0.939) has a slight correction for the number of predictors, however, a very high R-Squared indicating a very good model fit. The regression analysis of variance (ANOVA) shows F to have 263.24 and Significant level of 0.000. The F-statistic is very high and sig. = 0.000, meaning the regression model is statistically significant overall. The predictors (Injured, Killed, and People Involved) collectively predict Cases very well.

Table 2: Coefficients of the Variables

Predictor	B (Unstandardized)	Beta (Standardized)	t-test	Significant
(Constant)	3.922	-	0.577	0.566
Injured	0.318	0.537	5.108	0.000
Killed	0.701	0.366	3.859	0.000
People_Involved	0.044	0.285	2.710	0.010

We observed that all three predictors are statistically significant with ($p < 0.05$). Injured has the strongest impact ($\text{Beta} = 0.537$) to predicting accident cases; Killed also has a strong positive effect and followed by “People Involved” contributes moderately but significantly. The study shows that for each additional person killed, cases increase by approximately 0.70; and for extra person involved, cases increase by 0.044.

Based on our model, the final regression equation for predicting future occurrences is written as:

$$\text{Cases} = 3.922 + 0.318(\text{Injured}) + 0.701(\text{Killed}) + 0.044(\text{People_Involved}).$$

Residual Plot

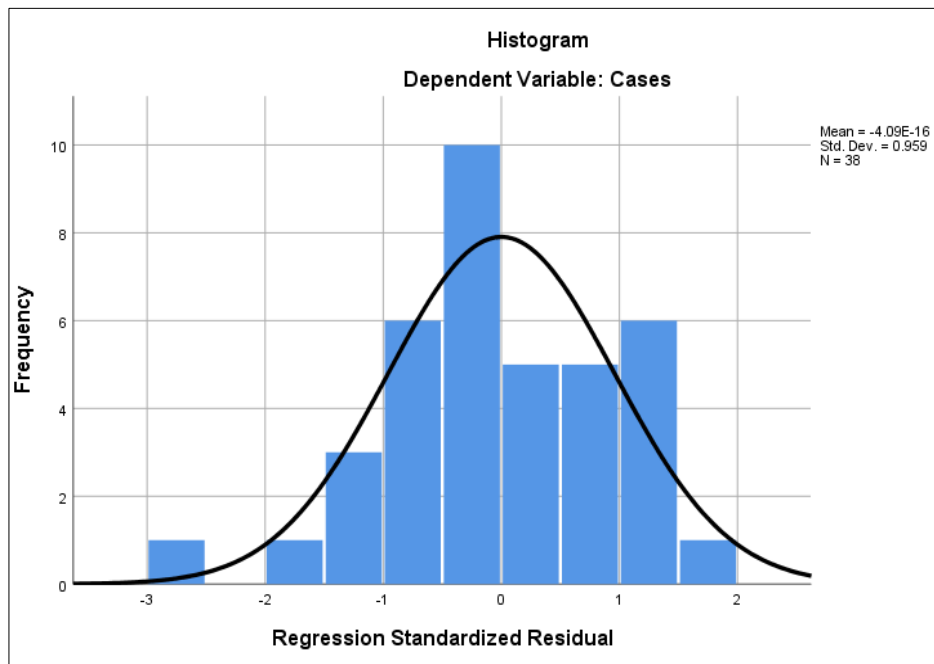


Fig 2: Residual plot of the variable without squared terms

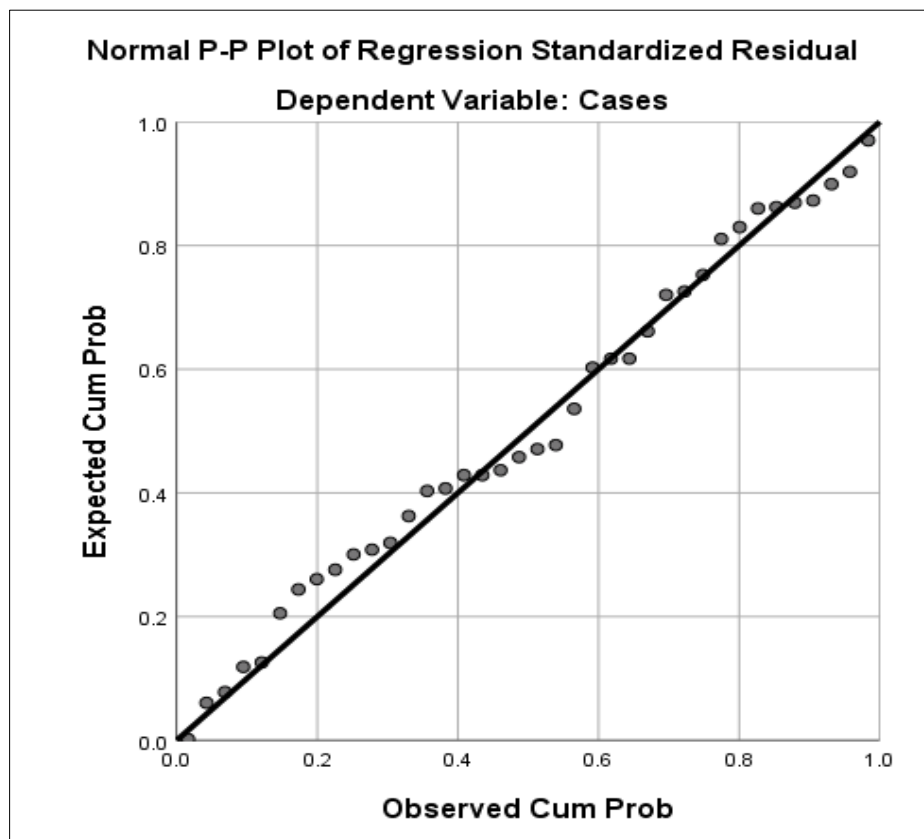


Fig 3: N-N plot of regression standardized residual

We observed the clustering of the data around the center line. This is good and it shows that most residuals are close to zero, meaning the model's predictions are fairly accurate. However,

there is a wiggle of the data which could indicate a mild non-linearity or a pattern the linear model is not fully capturing at extreme values. This could also suggest heteroscedasticity,

especially if the curve involves the spread changing (widening or narrowing). Since the curve happens almost twice or thrice and very close to the center and at the end, it might reflect outliers or leverage points. This may likely be showing extreme values within quarter four of 2023 or 2019. Although, it is not immediate problem since the curve is small and the overall scatter is fairly random.

Possible presence of curvature called for the running of a non-linearity test (e.g., adding squared terms to the model). With that singular action, appropriate normality of results was revealed as shown below.

From the scatter plot, we observed few outlier points at extreme tail end of the plot. Partial plot also reveals an outlier point in the data. The scatter plot shows the randomness of the data to be good.

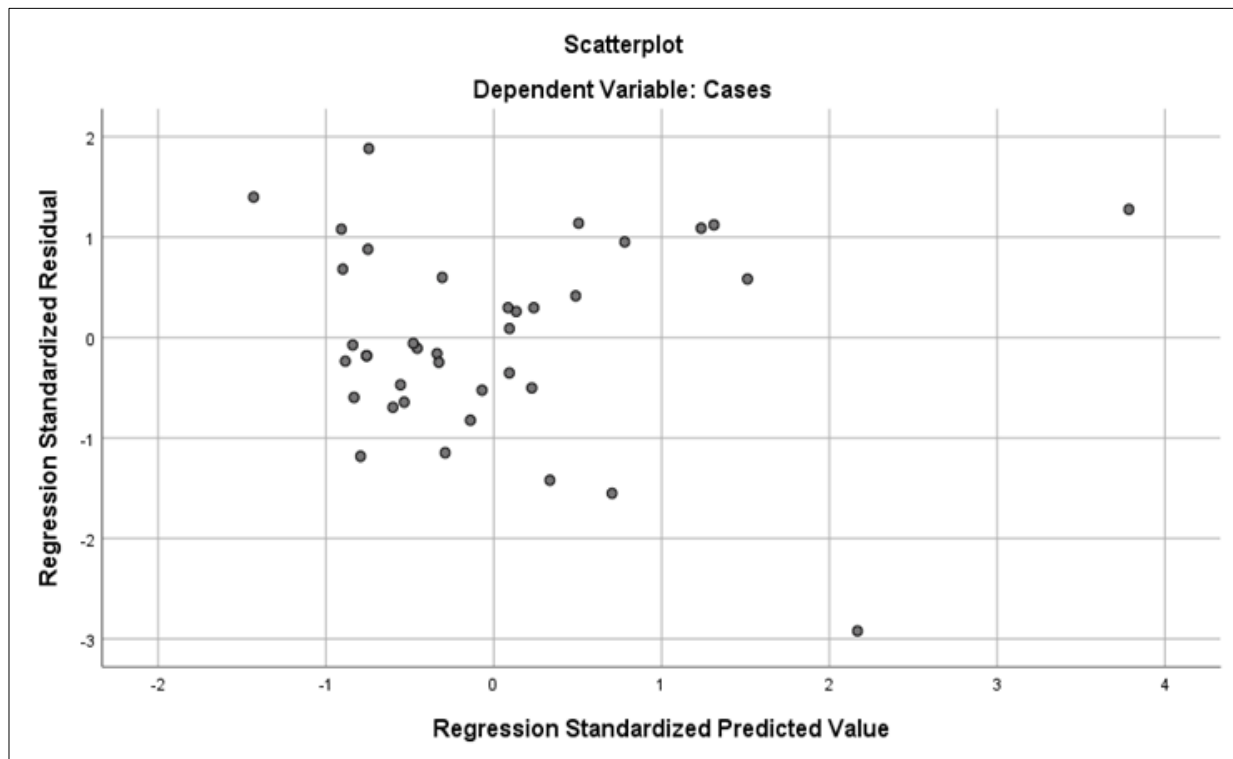


Fig 4: Scatter plot showing accident cases.

Non-linearity Test Results (Addition of Squared Terms)

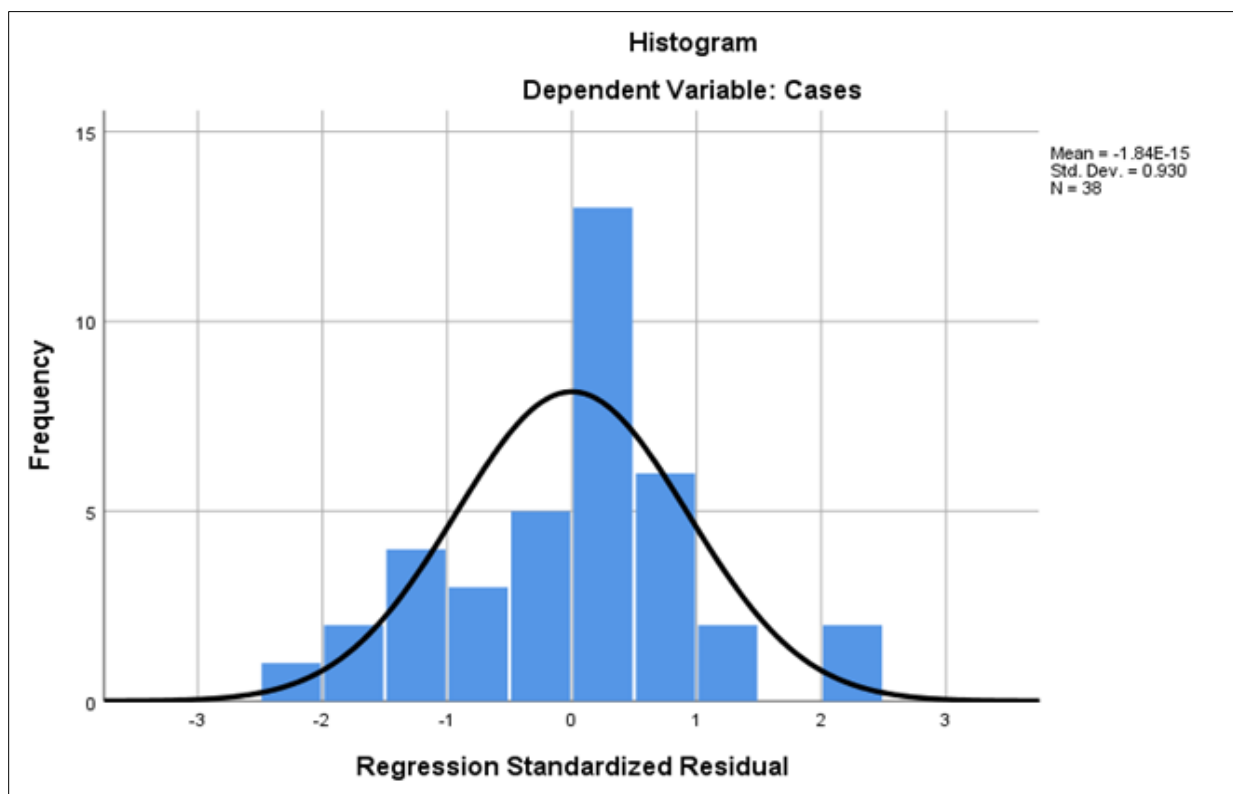


Fig 5: Histogram showing a normal distribution of variable

The histogram plot indicates normally distributed variables due to introduction of squared-terms to adjust for non-linearity in the dataset.

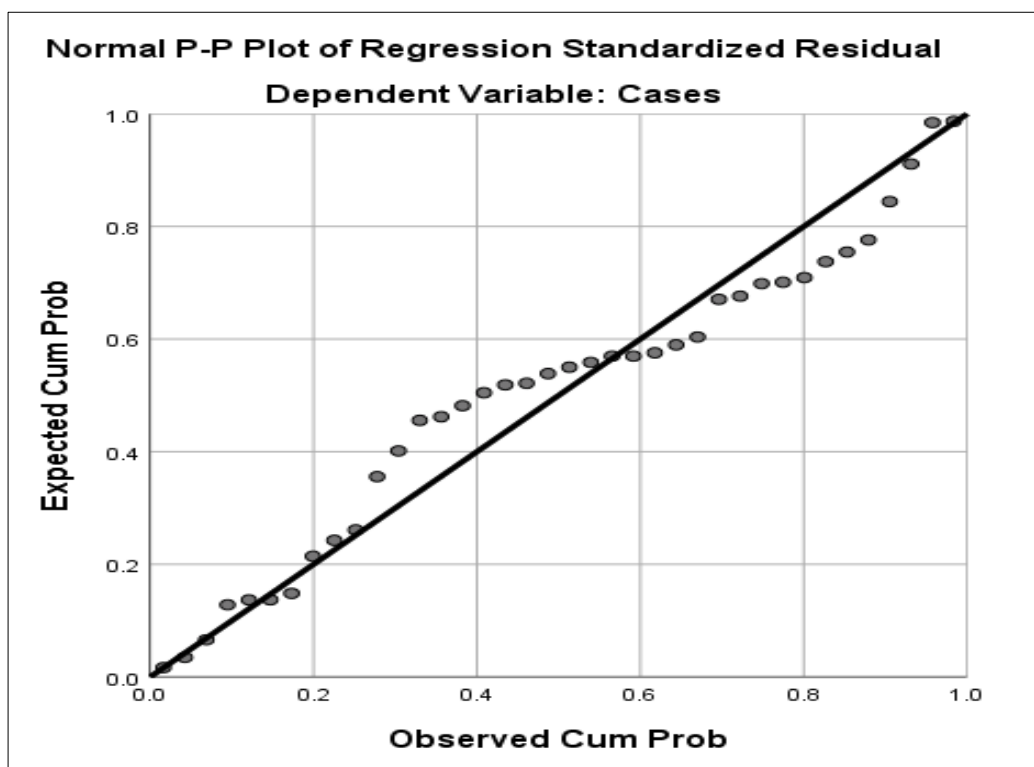


Fig 6: Normal P-P regression standardized residual

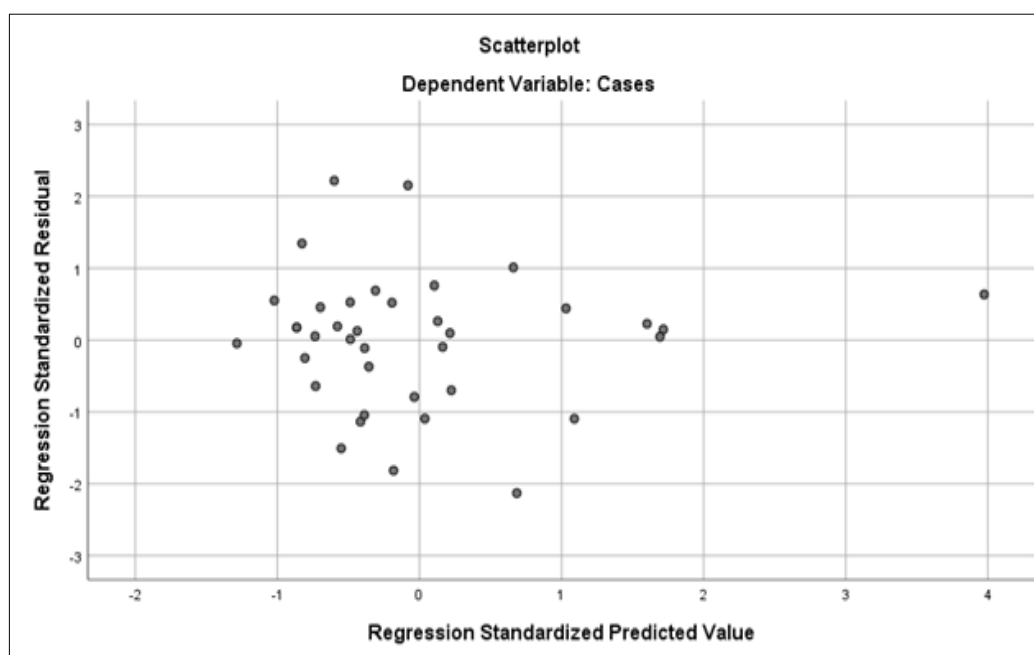


Fig 7: Scatterplot of standardized residual

3.4 Diagnostic Tests

Table 3: Model Summary

Model Summary ^b									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.907 ^a	0.822	0.807	15.794	0.822	52.465	3	34	0.000
a. Predictors: (Constant), People_Involved, Killed, Injured									
b. Dependent Variable: Cases									

From the diagnostic test, R-Squared recorded 0.822 which still suggest the model fit is good. However, the multicollinearity present (VIF > 8), tolerance value < 0.1 signal a problem and this might affect reliability of individual coefficient. Notwithstanding, all predictors are significant and each contributes to predicting cases, hence, predictors are reliable.

By the time Killed-Square terms were added into the model, the results show:

1. Original Model that has R-Squared 0.822
2. With Killed - Square, value of R^2 increased from 0.822 to 0.846.
3. That means R^2 increased from 82.2% to 84.6%; indicating the model with a non-linear term explains more variance in accident cases. The curve observed above in residuals is partially accounted for by the Squared terms as shown.

Killed-Squared is statistically significant confirming existence of Non-linear relationship at higher levels of **fatalities**, cases rise faster than linearly (e.g. sharp jumps in extreme data quarters like last quarters of 2019 - and 2023 respectively).

3.4.1 Multicollinearity Check (VIF)

The Variance Information Factor (VIF) “injured” recorded 9.46, this shows high multicollinearity, killed had 3.92 indicating acceptability and people involved recorded 8.62, also high multicollinearity; while constant with 6.16 is not a concern for multicollinearity.

The detection of Multicollinearity in the data informed the needs to consider Principal Component Analysis approach to eliminate the correlated explanatory variables. PCA helps the researcher to Standardized the predictors (mean =0, standard deviation (SD) =1). It enables the researcher also to create new uncorrelated variables called principal components and these components captures maximum variance in the study predictors. We observed that PCA reduces multicollinearity among predictors, transform correlated variables (like Injured, Killed, People involved) into uncorrelated components and helps in dimensionality reduction while preserving most of the data's variance.

3.4.2 Standardization Formula

Before we applied PCA, all variables were standardized to have mean =0 and standard deviation =1, using the formula below.

$$Z = \frac{X - \bar{X}}{\sigma}$$

Where,

X = Original value

$\bar{X}$ = Mean of the variables;

σ = Standard deviation and

Z = Standardized value

3.4.3 Covariance Matrix Calculation

Covariance matrix calculation after standardization, we used the covariance matrix of the standardized variables as stated below.

$$C = \frac{1}{n-1} Z^T Z$$

Where;

C = Covariance matrix, Z^T = Transpose of the standardized data matrix,

n = Number of observations

3.4.4 Eigen Decomposition

To extract the principal components, we solved the eigenvalue equation.

$$CV_i = \lambda V_i$$

Where;

λ_i = eigenvalue amount of the variance explained by component i.

V_i = eigenvector (direction of component i)

3.4.5 Principal Component Score Formula

We calculated each principal component using the formula below as:

$$PC_i = Z_1.a_1 + Z_2.a_2 + Z_3.a_3 + \dots + Z_p.a_p$$

Where:

Z = Standardized data, V= Matrix of eigenvectors, PC= Principal component scores

a_j = Loading of variable j on component i

We applied PCA to the standardized predictors: “Injured, Killed, and People-Involved” and the results below were obtained.

Table 4: Correlation Matrix

		Injured	People Involved
Correlation	Injured	1	0.937
	Killed	0.856	0.841
	People_Involved	0.937	1

Table 5: PCA Extraction Method

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.757	91.894	91.894	2.757	91.894	91.894
2	0.181	6.038	97.932			
3	0.062	2.068	100			

The above table reveals that PC1 alone captures almost all the information (91.9%) in the three original predictors. While PC2 adds only 6.0%, meaning together PC1 AND PC2 represent 98%. This confirms that the predictors are strongly

correlated, which principal component analysis has resolved by creating uncorrelated axes.

Table 6: Regression Model Analysis Using PC1 and PC2

Model Type	R - Squared
Original Regression	0.822
PCA Regression (PC1 + PC2)	0.822

From the table 3.4.7 above, the PCA-based model explains 82.2% of the variance in cases identical to the original model using injured, killed, and people-involved. This shows that PCA successfully retained predictive power while eliminating

multicollinearity. PC1 and PC2 are orthogonal, that is, uncorrelated.

Regression Model

In regression models, we run two different models for the analysis.

Regression Model 1: School Dropout Model

Predicting school dropout from accident count, people involved, casualties, where casualties is the sum of people injured and killed.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$\text{SchoolDroppedOut} = \beta_0 + \beta_1(\text{AccidentCounts}) + \beta_2(\text{PeopleInvolved}) + \beta_3(\text{Casualties}) + \varepsilon$$

Where:

X_1 is persons injured

X_2 is persons killed or died

X_3 : people involved and

Y_1 : School Dropped out (Binary)

Y_2 : Family Income (Continuous, Naira in thousands)

Y_3 : Job Loss (Binary)

The result of the logistic regression is state below:

Table 7: Coefficients

Predictor	B (Coefficients)	Sig. (p)
Accidents	0.021	0.029
Casualties	0.017	0.041
People Involved	0.003	0.457

$$\text{FamilyIncome} = \beta_0 + \beta_1(\text{AccidentCounts}) + \beta_2(\text{PeopleInvolved}) + \beta_3(\text{Casualties}) + \varepsilon$$

The multiple regression analysis reveals the following results as presented in a table 8 below:

Table 8: Multiple Regression Analysis

Predictor	B (Coefficients)	Sig. (p)
Accidents	-0.551	0.021
Casualties	-0.312	0.136
People Involved	-0.122	0.176

In this model, we noted that a 1-unit increase in accidents is associated with a N551,000 decrease in family income, and is statistically significant. However, "Casualties" have negative impact on income, but not statistically significant. Also, the people involved value is negative but not significant. This study indicates an increase in road accidents which significantly reduces household income, likely due to medical costs, job loss, or vehicle loss. The dependent variable (e.g. dropout) increases as the independent variable increases.

4. Discussion of Findings

The data support the hypothesis that accident frequency and severity negatively influence income, education and invariably the health of the members of such households. Regression diagnostics confirm that assumptions were adequately met. Logistic regressions show that even moderate increases in accident figures correspond to rising dropout likelihood. Though job loss results were not statistically significant, patterns suggest directional alignment.

A 1-unit increase in accidents is associated with a 2.1% increase in school drop cases, holding other variables constant. This is statistically significant. In case of casualties, we observed a 1-unit increase in casualties leads to an estimated 1.7% increase in dropout, also significant. The people involved is not statistically significant; changes in people involved don't meaningfully affect dropout in this model. Road accidents and casualties positively and significantly increase school dropout rates in Delta State.

Regression Model 2: Predicting Family Income from same predictors

There is a very strong positive correlation ($r = 0.971$) between observed and predicted values of Cases. The researcher observed also that all three predictors are statistically significant with ($p < 0.05$). Injured has the strongest impact (Beta = 0.537) to predicting accident cases; Killed also has a strong positive effect and followed by "People Involved" contributes moderately but significantly. The study shows that for each additional person killed, cases increase by approximately 0.70; and for extra person involved, cases increase by 0.044.

Findings revealed that the families affected by these accidents often face immediate financial crises, particularly when victims are primary earners. The ripple effects include disrupted education for children, increased household debt, and strained community networks.

The outcome of the regression modelling supports Aluko *et al.* (2019)^[6] study on impact of RTAs on productivity loss in Southwest Nigeria, that revealed a strong relationship between accident involvement and loss of income. The negative effects of the accidents on families in terms of children's education, feeding and reduction in social lives in Delta State are in consonant with the FRSC (2021) report.

World Health Organization (WHO, 2022) stated that road injuries are among the top 10 causes of death in Nigeria, highlighting the need for data-driven policy. The findings agreed with the WHO that there is a strong positive correlation between road accidents and injuries. Existing safety strategies to curb road traffic accidents be intensified to reduce accidents in Delta State, the hub of Nigeria Oil production and extraction. FRSC should endeavour to maintain roads in Delta cities, enforce speed limit policy,

check unwholesome behaviours among road users, and violators should face the wrath of the law to reduce accidents to the barest minimum in the state.

The 2019 witnessed the peak period of RTAs under the period of our investigation which tallied with COVID-19 pandemic period. This could have been increased in travelling due to fear of COVID-19 pandemic in urban areas and people travelled to their villages for safety.

Modeling the socio-economic impact indicates a strong positive correlation between road traffic accidents and people injured, killed, and people involved. However, we noticed some of the figures provided lack integrity either due to under reporting, over reporting, entering data errors or intentional falsification of data by the collectors. Therefore, there is a need to promote good data-driven policy orientation within the data collectors and documentation officials, to improve on the quality of data at all time.

The scatter plot derived from regression modeling indicates a strong positive relationship between the first principal component (PC1) and accident cases, supporting the result that PC1 is a strong predictor. The PC1 holds most of the predictive power. PC2 explains minor variation and may not be needed for practical modeling.

The model stability was sensitive to overlap in the predictors before introduction of PCA, but more stable under collinearity after PCA's application. The application of PCA, changed correlated variables to uncorrelated (orthogonal) components, solving the multicollinearity problem, hence, improved the reliability and stability of the data. Therefore, data on road traffic accidents should ensure principal component analysis is always carried out during analysis to enhance data quality and reporting.

5.1 Summary of Findings

This study modeled the effects of road traffic accidents on school dropout, job loss, and family income using multiple regression and PCA. Accident-related variables were shown to significantly predict social and economic disruptions in Delta State. The scatter plot shows a strong positive relationship between the first principal component (PC1) and accident Cases, supporting the result that PC1 is a strong predictor. The adoption of PCA in the analysis aids the reduction of the dimensionality of the data by transforming correlated variables into a smaller set of uncorrelated principal components. These components represent linear combinations of the original variables, capturing the most significant sources of variation in the data. There is a strong positive relationship between road traffic accident counts and the explanatory variables (injured, killed and people involved).

5.2 Conclusion

The study concludes that road traffic accidents have a measurable and statistically significant impact on the social and economic well-being of individuals and households in Delta State. This is because an increase in road traffic accidents significantly reduces household income, likely due to medical costs, job loss, or vehicle loss. The dependent variable (e.g. dropout) increases as the independent variable increases which was represented by positive Beta (β) in the study, while negative β , shows the dependent variable (e.g. income) decreases as the independent variable increases. This finding confirms the need for urgent data-driven policy in monitoring and evaluation of roads usage in Delta Cities.

5.3 Recommendations

This comprehensive statistical analysis reinforces the role of RTAs in worsening socioeconomic conditions. Hence, we recommend that the authority saddled with road safety management should:

- Strengthen road safety education, drivers training, and enforcement to reduce accident rates.
- Improve post-accident support and social welfare systems to cushion economic losses.
- Use data analytics in policymaking to identify accident hotspots and target interventions.

5.4 Suggestions for Further Research

Future studies can apply panel data techniques or include other States for comparative analysis. Primary data collection could offer deeper insights into household-level experiences, hence provides a robust statistical analysis which could reveals more hidden effects on the families, communities, and State respectively.

5.5 Contribution to Knowledge

This study demonstrates how secondary data and statistical modeling can provide robust evidence for social policy in road safety and accident impact mitigation. It also provides a statistical foundation for understanding how road traffic accidents affect family units in Nigeria, especially Delta State context. By integrating time-series trends and multivariate modeling, the research contributes:

- A replicable framework for analyzing accident trends and consequences
- A model for predicting family distress based on accident severity and socioeconomic factors
- Empirical evidence to guide social welfare policy and transport safety planning.

References

- Adebayo SO. Distracted driving and its implications on road safety in Southern Nigeria. *African Journal of Transportation Studies*. 2020;14(2):44-56.
- Adebayo SA, Nwachukwu CE. Enforcement intensity and accident mortality in Nigeria. *Journal of African Transport*. 2020;8(2):101-115.
- Adeniji K, Ojo AA. Trends in road traffic accidents in Nigeria (2006-2016): an empirical analysis. *Nigerian Journal of Statistics*. 2018;30(1):59-70.
- Adesanya A, Odetunde J. Socio-economic costs of road accidents on households in South-West Nigeria. *Journal of Social Policy Studies*. 2020;11(1):35-49.
- Adedotun T, Okon B, Ibrahim A, Eze C. Psychosocial consequences of road traffic accidents on Nigerian households. *Journal of Public Health and Social Policy*. 2024;18(2):45-62.
- Aluko OO, Adebayo AM, Adebisi TF, Abiola AO, Oni OF. The burden of road traffic crashes in Southwest Nigeria: a literature-based analysis of productivity loss. *African Journal of Health Economics*. 2019;8(1):1-10. <https://doi.org/10.35202/AJHE.2019.8110>
- Atubi AO. Road traffic accidents occurrences in Lagos State: causes and preventive measures. *Journal of Sustainable Development in Africa*. 2012;14(1):25-35.
- Atubi AO. Determinants of road accident occurrences in Lagos State: some lessons for Nigeria. *International Journal of Humanities and Social Science*. 2012;2(6):252-259.

9. Eze P, Nwanekezie C. Psychological and economic effects of road accident fatalities on families in Enugu State, Nigeria. *Nigerian Journal of Health and Social Research*. 2018;7(3):23-39.
10. Federal Road Safety Corps (FRSC). Annual Road Traffic Crash Report. Abuja: FRSC Publications; 2021.
11. Federal Road Safety Corps (FRSC). Annual Road Traffic Crash Report. Abuja: FRSC Publications; 2024.
12. Gujarati DN, Porter DC. Basic Econometrics. 5th ed. New York: McGraw-Hill; 2009.
13. Ogunbodede EF. Urban road transportation in Nigeria from 1960 to 2006: problems, prospects and challenges. *Ethiopian Journal of Environmental Studies and Management*. 2008;1(1):7-17.
14. Okafor CU, Eke JO. Regression analysis of road traffic accident fatalities in Enugu State, Nigeria. *International Journal of Scientific and Engineering Research*. 2017;8(6):1428-1435.
15. Okonkwo MJ, Adebajo TA. Road traffic crashes in South Nigeria: a spatial-temporal analysis. *African Journal of Safety Science*. 2019;4(1):77-92.
16. Oluwasanmi O, Ibrahim T, Bello A. Temporal trends in urban road accidents: evidence from Lagos. *Nigerian Statistical Review*. 2021;19(3):55-72.
17. Olagunju TA, Udo BC. A statistical approach to causes of road traffic crashes in Nigeria. *International Journal of Statistical Research*. 2019;8(1):14-21.
18. Oni SI, Okanlawon KR, Adeyemo OT. Time-series analysis of road traffic accident patterns in Lagos State, Nigeria. *Journal of Emerging Trends in Economics and Management Sciences*. 2011;2(4):288-292.
19. Uzundu CA, Nduka KI. Analysis of road traffic accidents in Nigeria's South-South region: implication for policy and planning. *International Journal of Transport and Logistics*. 2020;4(2):45-56.
20. World Health Organization (WHO). Global status report on road safety. Geneva: WHO; 2023.