Journal of the Nigerian Association of Mathematical Physics
Volume35,(May, 2016), pp193 – 200
© J. of NAMP

A Multivariate Linear Regression Theoretic Approach to Modelling Road Traffic Accidents

B.U Oreko[1], S. Okiy[3],C.C Nwobi-Okoye[2] and A.C Igboanugo [4]

[1]Department of Mechanical Engineering, Federal University of Petroleum Resources, Effurun, Nigeria.
[2]Department of Welding Engineering and offshore Technology, Petroleum Training Institute, Effurun, Nigeria
[3]Anambra State University (Chukwuemeka Odumegwu Ojukwu University), Uli, Nigeria
[4]Department of Production Engineering, University of Benin, Benin City, Edo State, Nigeria.

Abstract

This work investigated the relationship that exist among road traffic accident outcomes in Nigeria namely; total road traffic cases, fatal accident cases, serious accident cases, minor accident cases, number of persons killed, number of persons injured and total casualty. A historical data of fifty-four years period (1960-2013) were obtained from the Department of Policy, Research and Statistics, Federal Road safety corps FRSC, Nigeria. Multivariate linear regression technique was adopted to explore the interrelatedness that exists among these road traffic accident outcomes in Nigeria. SPSS, MATLAB AND EXCEL 2010 were used for the analysis. Each of these variables was used as response variable while others variables become the predictors. It was found that at 95 percent confident interval, Multicollinearity exist between total accident cases, fatal accident cases, serious accident cases, minor accident cases. Also strong correlation exists between total casualty and the number of persons killed and number of person injured. However, weak correlation was found to exist as the response variable is replaced interchangeably. From the result it was seen that every accident case is likely to be either serious or minor accident case. Hence concerted efforts are needed by all road users to comply with road traffic regulations.

Keywords: Road traffic cases, fatal accident cases, serious accident cases, minor accident cases

1.0 Introduction

The relationship among road traffic accident outcomes appears not to be well understood for example, if accident happens, what is the likelihood that it becomes fatal, serious, minor, persons were killed, persons were injured or casualty exist. It is suspected that a causal relationship exist among these variables. And that the accidents causation process tends to have a pattern which can be reproduced by obtaining sample realization. A good understanding of these variables will not only help the policy makers in decision making but will also encourage road users and all stakeholders in complying with traffic rules and regulations. Knowing that each time there is a road traffic accident, the probability exists of varied ranges of losses, which are not in anybody interest. A handful of literature exists on road traffic accidents. Venkataraman et al [1] studied the frequency of occurrence of highway accidents on the basis of a multivariate analysis of roadway geometrics (e.g. horizontal and vertical alignments), weather, and other seasonal effects. The results of their analysis gave insight into potential measures to counter the adverse effects of weather on highway sections with challenging geometrics. Karlaflis and Golias [2] analyzed the relationship between rural road geometric characteristics, accident rates and prediction and observed that geometric design variables and pavement condition variable are the most important factors affecting accident rates. Relatedly, other researchers [3-5] studied road traffic accidents in different scenarios using multiple linear and Poisson regression in their analysis. They established a relationship between mean absolute curvature and mean absolute grade alignment as important variables for accidents rate evaluation. Also multiple and Poisson regression were used in modelling various accident parameters such as

Corresponding Author:B.U Oreko, E-mail:orekoben@yahoo.com, Tel.: +2348036785421

the mortality rates and number of deaths in a specific population, the relation of demographic, geographical, and temporal explanatory variables with mortality in difference regions, compare age- and sex-specific mortality rates due to injuries [6, 7]. Other statistical method such as Log-linear regression, logarithmic transformation and spatial autocorrelation corrected regression model were employed to study the effect of age, period of death and birth in motor vehicle mortality, demographic and environmental, pedestrian injuries and fatal rate [8-11]. Furthermore, multivariate linear regression has also been applied in the analysis of road traffic accidents in some regional areas in Nigeria. Atubi [12] applied multiple regression analysis to determine the proportion of variation of road traffic accidents as dependent variable to length of roads, presence of road safety and population as independent variables in Lagos state between 1970 and 2001, and found that accidents increase with increased length of roads, presence of road safety corps and population in Lagos state. Also several scholars [13-15], employed multiple linear regressions to road traffic accident characteristics. From the forgoing review, it is evident that literature abound on road traffic accidents variables with respect to various factors ranging from road characteristics to human error. However from available literature that exist, the interrelatedness among road traffic accident variables such as fatal accidents, serious accidents, minor accidents, number of person killed, number of persons injured and accident casualties has rarely been explored especially in Nigeria. Hence this study seeks to analyze the interrelationship between road accident cases, fatal accidents, serious accidents, minor accidents, number of person killed, number of persons injured and accident casualties using multivariate linear regression modelling approach.

2.0 Materials and Methods

Road traffic Accidents data covering a fifty-four (54) years period (1960-2013) were obtained from the department of Policy, Research and Statistics, Federal Road Safety Corps of Nigeria (FRSC). The data were categorized into seven road traffic accident outcomes namely; Total accident cases, Fatal accident cases, serious accident cases, minor accident cases, number of persons killed, number of person injured and total casualty. According to FRSC, these variables may be defined as follows;
Fatal Case: A case where at least one death is recorded
Serious Cases: A case where person(s) is/are injured but no death
Minor Case: When no death or injury is recorded
Total Cases = Fatal + Serious + Minor cases
Total Casualty = Persons killed + persons injured
Let
$Y = Y_{\text{Total accident cases}}$ = Total accident cases
$X_1 = X_{\text{Fatal cases}}$= Fatal accident cases
$X_2 = X_{\text{Serious cases}}$= Serious accident cases
$X_3 = X_{\text{Minor cases}}$ = Minor accident cases
$X_4= X_{\text{No. killed}}$= Number of persons killed
$X_5 = X_{\text{No. injured}}$ = Number of persons injured
$X_6 = X_{\text{Total casualty}}$ = Total accident casualty
The general form of a multiple regression model can be expressed as

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots.. +\beta_nX_n+\varepsilon \quad (1)$$

where
Y = dependent variable
X = independent or predictor variable
β = coefficients
ε = residual error

From equation (2) we obtained the normal equations, and the corresponding regression parameters β_o β_1 …..β_n. Computations were manually done and validated using SPSS and MATLAB software package.

$$\sum Y = n\beta_0 + \beta_1\sum X_1 + \beta_2\sum X_2 + \beta_3\sum X_3 + \beta_4\sum X_4 + \beta_5 \sum X_5 +\beta_6\sum X_6 \quad (2)$$

$$\sum X_1Y = \beta_0\sum X_1+\beta_1\sum X^2_1 +\beta_2\sum X_1X_2 + \beta_3\sum X_1X_3 + \beta_4 \sum X_1X_4 + \beta_5\sum X_1X_5 + \beta_6 \sum X_1X_6 \quad (3)$$

$$\sum X_2Y= \beta_0\sum X_2+ \beta_1\sum X_1X_2 + \beta_2\sum X^2_2 + \beta_3\sum X_2X_3 + \beta_4\sum X_2X_4 + \beta_5\sum X_2X_5 + \beta_6\sum X_2X_6 \quad (4)$$

$$\sum X_3Y = \beta_0\sum X_3 + \beta_1\sum X_1X_3 + \beta_2\sum X_2X_3 + \beta_3\sum X^2_3+ \beta_4\sum X_3X_4 + \beta_5\sum X_3X_5 + \beta_6 \sum X_1X_6 \quad (5)$$

$$\sum X_4Y = \beta_0\sum X_4 + \beta_1\sum X_1X_4 + \beta_2\sum X_2X_4 + \beta_3\sum X_3X_4 + \beta_4\sum X^2_4 + \beta_5\sum X_4X_5 + \beta_5\sum X_4X_6 \quad (6)$$

$$\sum X_5Y = \beta_0\sum X_5 + \beta_1\sum X_1X_5+ \beta_2\sum X_2X_5 + \beta_3\sum X_3X_5 + \beta_4\sum X_4X_5 + \beta_5\sum X^2_5 + \beta_6\sum X_5X_6 \quad (7)$$

$$\sum X_6Y= \beta_0\sum X_6 + \beta_1\sum X_1X_6+ \beta_2\sum X_2X_6 + \beta_3\sum X_3X_6 + \beta_4\sum X_4X_6 + \beta_5\sum X_5X_6 + \beta_6\sum X^2_6 \quad (8)$$

And fitted model becomes;

$$\hat{Y} = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5 X_5 + \beta_6X_6 \quad (9)$$

Also the Coefficient of Determination can be expressed as, $R^2 = 1 - \frac{\sum(Y-\hat{Y})^2}{\sum(Y-\bar{Y})^2}$ (10)

Coefficient of Correlation, $R=\sqrt{1-\frac{\sum(Y-\hat{Y})^2}{\sum(Y-\bar{Y})^2}}$ (11)

Covariance; $COV(X,Y) = \frac{\sum(X_t-\bar{X})(Y_t-\bar{Y})}{n-1}$ (12)

Pearson correlation coefficient, $r_{xy} = \frac{cov(X,Y)}{S_x S_y}$ (13)

$S_x = \sqrt{\frac{\sum(X_t-\bar{x})^2}{n-1}}$ (14)

$S_y = \sqrt{\frac{\sum(Y_t-\bar{Y})^2}{n-1}}$ (15)

3.0 Results and Discussion

The relationship between total accident cases, fatal accident cases, serious accident cases, minor accident cases, number of persons killed, number of persons injured were first developed. Subsequently, the other variables were used as dependent variables, in order to determine the multicollinearity that exists among the variables.

i. The fitted model equation to predict the relationship between total accident cases (dependent variable) fatal cases, serious cases, number of persons killed and number of persons injured (independent variables) is given as;

$\hat{Y}_{\text{Total accident cases}} = -2130 + 0.370 (X_{\text{fatal cases}}) + 2.531(X_{\text{serious cases}}) + 0.554 (X_{\text{no.killed}}) - 0.232 (X_{\text{no. injured}})$ (16)

ii. The fitted model equation to predict the relationship between fatal cases (dependent variable), total accident cases, serious cases, number of persons killed and number of persons injured (independent variables) is given as;

$\hat{Y}_{\text{Fatal cases}} = 2010 + 0.056 (X_{\text{total accident cases}}) - 0.365 (X_{\text{serious cases}}) + 0.649 (X_{\text{no. killed}}) - 0.030 (X_{\text{no.injured}})$ (17)

iii. The fitted model equation to predict the relationship between serious cases (dependent variable), total accident cases, fatal cases, number of persons killed and number of persons injured (independent variables) is given as;

$\hat{Y}_{\text{Serious cases}} = 1611 + 0.337(X_{\text{total accident cases}}) - 0.321(X_{\text{fatal cases}}) - 0.004 (X_{\text{no. killed}}) + 0.074 (X_{\text{no. injured}})$ (18)

vi. The fitted model equation to predict the relationship between minor cases (dependent variable), total accident cases, fatal cases, number of persons killed and casualty (independent variables) is given as;

$\hat{Y}_{\text{Minor cases}} = -1611 + 0.663 (X_{\text{total accident cases}}) - 0.679 (X_{\text{fatal cases}}) + 0.078 (X_{\text{no. killed}}) - 0.074 (X_{\text{casualty}})$ (19)

v. The fitted model equation to predict the relationship between number of persons killed (dependent variable), total cases, fatal cases, number of persons injured and minor cases (independent variables) is given as;

$\hat{Y}_{\text{No. killed}} = -1671 + .115(X_{\text{total cases}}) + 0.939(X_{\text{fatal cases}}) + 0.116(X_{\text{injured cases}}) + 0.006 (X_{\text{minor cases}})$ (20)

vi. The fitted model equation to predict the relationship between number of persons injured (dependent variable), total accident cases, number of persons killed, fatal cases and minor cases (independent variables) is given as;

$\hat{Y}_{\text{No. injured}} = 5254+1.724 (X_{\text{total cases}}) + 2.817 (X_{\text{no. killed}}) -3.992(X_{\text{fatal cases}}) - 2.952 (X_{\text{minor cases}})$ (21)

vii. The fitted model equation to predict the relationship between total casualty (dependent variable) and fatal cases , minor cases, number of persons injured and total accident cases (independent variables) is given as;

$\hat{Y}_{\text{casualty}} = -1670 + 1.116 (X_{\text{injured cases}})+ 0.939(X_{\text{fatal cases}}) - 0.115(X_{\text{total accident cases}}) + 0.006(X_{\text{minor cases}})$ (22)

Table 1 shows the correlation that exists between total accident cases, fatal, serious, minor accident number of persons killed and number of persons injured. It was observed that strong correlation exist between accident cases and serious cases and minor cases.

Table 1: Summary of Pearson Correlations Matrix for Road Traffic Accident Variables (1960-2013)

Accident Variables	Total Accident Cases	Fatal Cases	Serious Cases	Minor Cases	Number Injured	Number Killed	Casualty
Total Cases	1.000	0.260	0.936	0.948	0.590	0.299	0.402
Fatal Cases	0.260	1.000	0.120	0.000	0.787	0.357	0.504
Serious Cases	0.936	0.120	1.000	0.860	0.525	0.411	0.474
Minor Cases	0.948	.000	0.860	1.000	0.360	0.113	0.190
Number Injured	0.590	0.787	0.525	0.360	1.000	0.641	0.790
Number Killed	0.299	0.357	0.411	0.113	0.641	1.000	0.977
Casualty	0.402	0.504	0.474	0.190	0.790	0.977	1.000

Table 4.2: Descriptive Statistics for Road Accident Variables (1960-2013)

	Range	Min.	Max.	Sum	Mean		Std. Deviation	Variance
ACCIDENT VARIABLES	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Total Cases	29400	8477	37881	$1.09x10^6$	$2.02x10^4$	$1.08x10^3$	7965	$6.34x10^7$
Fatal Cases	6793	193	6986	$1.78x10^5$	$3.304x10^3$	$2.35x10^2$	1724	$2.97x10^6$
Serious Cases	13200	4143	17352	$4.80x10^5$	$8.89x10^3$	$3.95x10^2$	2899	$8.401x10^6$
Minor Cases	18700	964	19624	$4.34x10^5$	$8.031x10^3$	$6.89x10^2$	5064	$2.564x10^7$
Number Killed	10300	1083	11382	$3.34x10^5$	$6.194x10^3$	$3.77x10^2$	2777	$7.71x10^6$
Number Injured	33400	7771	41165	$1.12x10^6$	$2.071x10^4$	$1.08x10^3$	7999	$6.399x10^7$
Casualty	37900	9303	47219	$1.45x10^6$	$2.687x10^4$	$1.36x10^3$	10009	$1.002x10^8$

The strong correlation between total accident cases and serious cases and total accident cases and minor cases shows that every accident case could likely be serious or minor. Generally, from the correlation matrix in Table 2, there is a decent degree of relationship among the RTA variables, which is an indicator that the variables could be predicted from one another using regression analysis. This further corroborates the work of Atubi [12] who predicted total accident cases in Lagos State Nigeria using some road traffic variables.

4.0 Conclusion

Being able to find the relationship among road traffic accident variables, as done in this work, is absolutely necessary. This is because the value of a variable could be predicted by knowing the values of other variable. Predicting the values of road traffic accident variables would be invaluable to government agencies and emergency services in accident preparedness and reduction of fatalities on our roads. As stated by Nwobi-Okoye [16], the government has to do the large chunk of work required to reduce road traffic accidents and its attendant consequences. But the government alone would not do it all, the citizens are expected to imbibe the culture of obeying basic road safety rules to complement government efforts in reducing road traffic accidents for a better society.

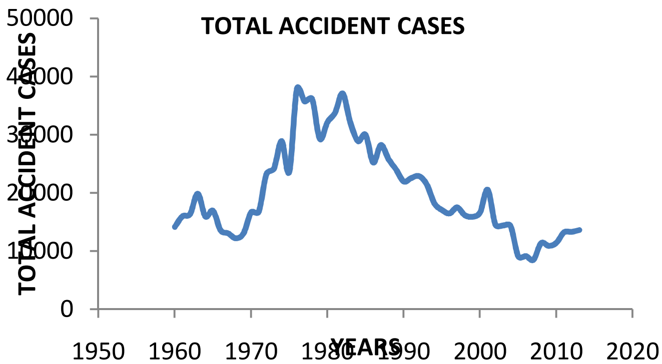
5.0 Appendix

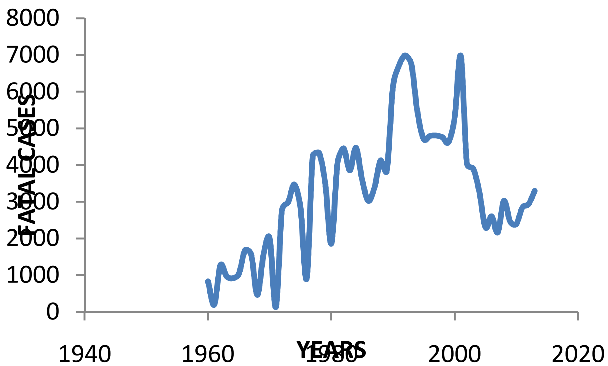
TABLE A1: SUMMARY OF PEARSON CORRELATIONS MATRIX FOR ROAD TRAFFIC ACCIDENTS IN NIGERIA (1960-1913)

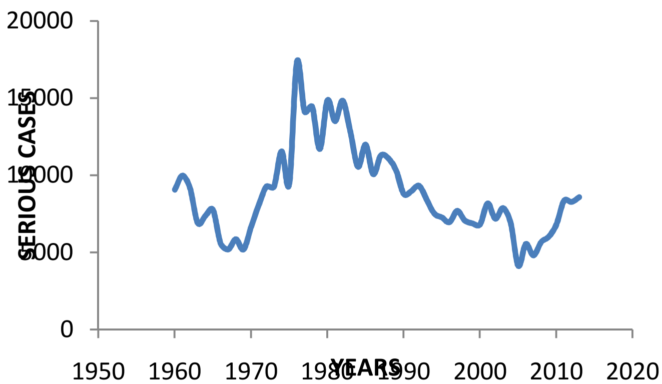
Accident Variables	Total Accident Cases	Fatal Cases	Serious Cases	Minor Cases	No. Injured	No. Killed	Casualty
Total Cases	1.000	0.244	0.940	0.950	0.580	0.298	0.399
Fatal Cases	0.244	1.000	0.120	0.550	0.787	0.357	0.504
Serious Cases	0.940	0.120	1.000	0.860	0.525	0.411	0.474
Minor Cases	0.950	0.550	0.860	1.000	0.360	0.113	0.190
No. Injured	0.580	0.787	0.525	0.360	1.000	0.641	0.790
No. Killed	0.298	0.357	0.411	0.113	0.641	1.000	0.977
Casualty	0.399	0.504	0.474	0.190	0.790	0.977	1.000

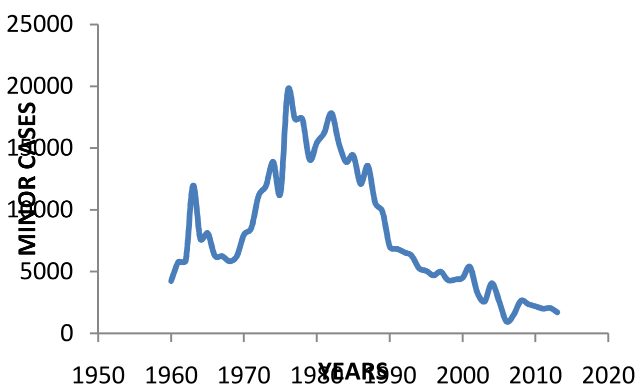
TABLE A2: DESCRIPTIVE STATISTICS FOR ROAD TRAFFIC ACCIDENTS IN NIGERIA (1960-2013)

Accident Variables	Sum Total	Mean	Std. Deviation	No. Of Period
Total Accident Cases	1091751	20217.6111	8086.43046	54
Fatal Cases	178406	3303.815	1723.59220	54
Serious Cases	479683	8883.0185	2898.50618	54
Minor Cases	433662	8030.7778	5063.49684	54
No. Killed	334472	6193.925	2776.92698	54
No. Injured	1116244	20671.185	7999.38500	54
Casualty	1450716	26865.11	10009.70458	54


TOTAL ACCIDENT CASES
TOTAL ACCIDENT CASES
YEARS
50000
40000
30000
20000
10000
0
1950
1960
1970
1980
1990
2000
2010
2020


FATAL CASES
YEARS
8000
7000
6000
5000
4000
3000
2000
1000
0
1940
1960
1980
2000
2020


SERIOUS CASES
YEARS
20000
15000
10000
5000
0
1950
1960
1970
1980
1990
2000
2010
2020


MINOR CASES
YEARS
25000
20000
15000
10000
5000
0
1950
1960
1970
1980
1990
2000
2010
2020

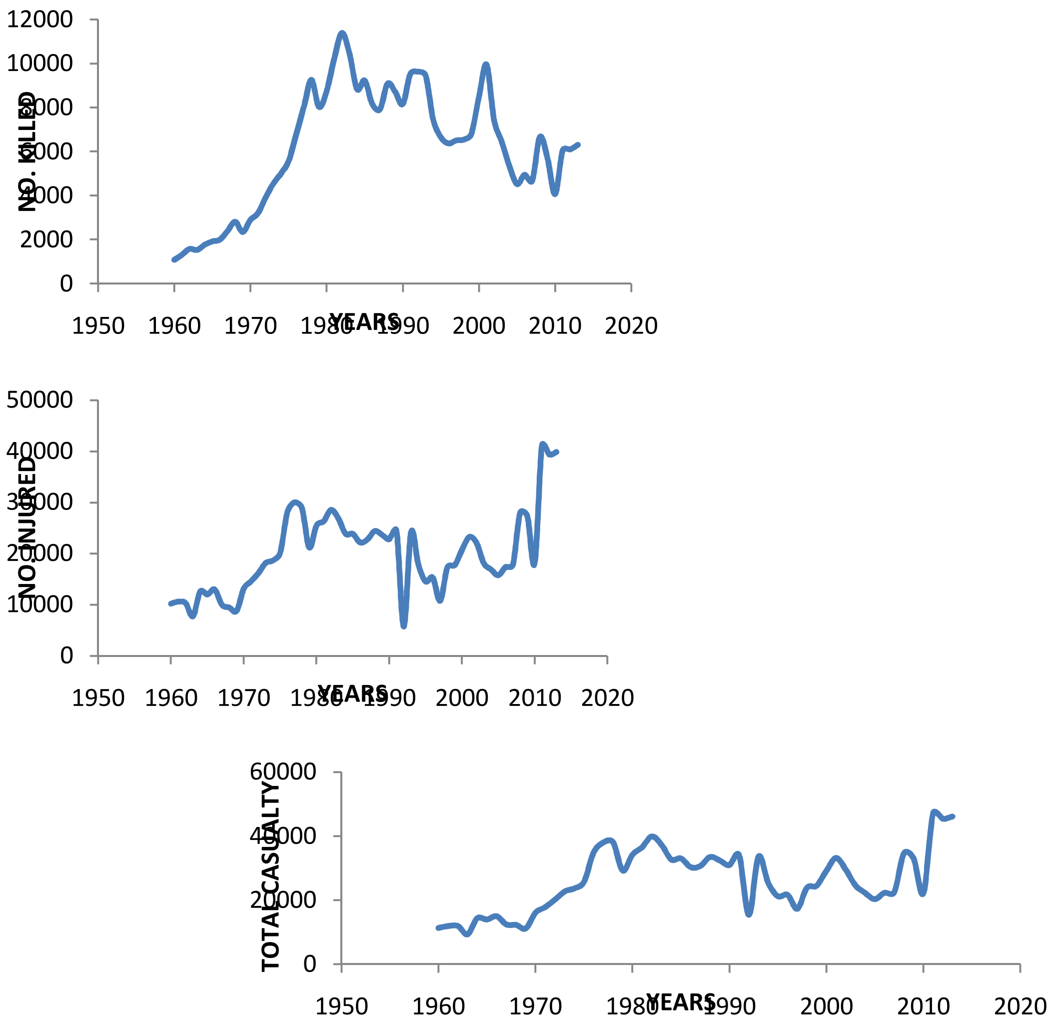


Figure A1. Graph of road traffic accident variables against number of years

TABLE A3: SUMMARY OF REPORTED ROAD TRAFFIC ACCIDENTS IN NIGERIA (1960-2013)

YEAR	FATAL CASES	SERIOUS CASES	MINOR CASES	TOTAL CASES	NO. PERSONS KILLED	NO. PERSONS INJURED	TOTAL CASAULTY
1960	826	9065	4239	14130	1083	10216	11299
1961	193	9982	5788	15963	1313	10614	11927
1962	1263	9159	5895	16317	1578	10341	11919
1963	967	6918	11950	19835	1532	7771	9303
1964	911	7371	7645	15927	1769	12581	14350
1965	1029	7762	8113	16904	1918	12024	13942
1966	1680	5600	6270	14000	2000	13000	15000
1967	1560	5200	6240	13000	2400	10000	12400
1968	459	5865	5839	12163	2808	9474	12282
1969	1559	5199	6230	12998	2347	8804	11151
1970	1999	6666	7991	16666	2893	13154	16047
1971	1129	8098	8518	17745	3206	14592	17798
1972	2782	9275	11130	23187	3921	16161	20082
1973	2981	9275	11925	24844	4537	18154	22691
1974	3467	11557	13869	28893	4992	18660	23652
1975	2834	9446	11331	23651	5552	20132	25684
1976	905	17352	19624	40881	6761	28155	34916
1977	4242	14140	17334	35351	8000	30023	38023
1978	4333	14444	17334	36111	9252	28854	38106
1979	3513	11708	14050	29271	8022	21203	29225
1980	1856	14855	15427	32138	8736	25484	34220
1981	4053	13510	16214	33777	10202	26337	36539
1982	4451	14838	17805	37094	11382	28539	39921
1983	3853	12844	15412	32109	10462	26866	37328
1984	4467	10557	13868	28892	8830	23861	32691
1985	3597	11991	14380	29978	9221	23853	33074
1986	3022	10075	12091	25188	8154	22176	30330
1987	3385	11286	13544	28215	7912	22747	30659
1988	4127	11091	10574	25792	9077	24413	33490
1989	3838	10314	9835	23987	8714	23687	32401
1990	6140	8796	6998	21934	8154	22786	30940
1991	6719	8982	6845	22546	9525	24508	34033
1992	6986	9324	6554	22864	9620	25759	35379
1993	6735	8443	6281	21459	9454	24146	33600
1994	5407	7522	5275	18204	7440	17938	25378
1995	4701	7276	5053	17030	6647	14561	21208
1996	4790	6964	4688	16442	6364	15290	21654
1997	4800	7701	4987	17488	6500	10786	17286
1998	4757	7081	4300	16138	6538	17341	23879
1999	4621	6888	4356	15865	6795	17728	24523
2000	5287	6820	4499	16606	8473	20677	29150
2001	6966	8185	5379	20530	9946	23249	33195
2002	4029	7190	3325	14544	7407	22112	29519
2003	3910	7882	2572	14364	6452	18116	24568
2004	3275	6948	4051	14274	5351	16897	22248
2005	2299	4143	2620	9062	4519	15779	20298
2006	2600	5550	964	9114	4944	17390	22334
2007	2162	4812	1503	8477	4673	17794	22467
2008	3024	5671	2646	11341	6661	27980	34641
2009	2460	6024	2370	10854	5693	27270	32963
2010	2388	6815	2182	11385	6052	35691	41743
2011	2840	8357	1999	13196	6054	41165	47219
2012	2935	8277	2050	13262	6092	39348	45440
2013	3294	8589	1700	13583	6544	40057	46601

Sources: Department of Policy, Research and Statistics, Federal Road safety Corps (FRSC) 2015.

6.0 References

[1] Venkataraman, F. M. and Woodrow B. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies Science Direct Accident Analysis & Prevention, Vol. 27, Issue 3, p 371–389
[2] Karlaftis, M.G. and Golias, I. (2002). Effect of road geometry and traffic volumes on rural roadway accident rate. Accident Analysis and Prevention 34, 357-365.
[3] Jones, I.S.and Whitfield, R.A. (1991). Predicting injury risk with new car assessment program crashworthiness rating. Accident Analysis and Prevention 6 (20), 411-419.
[4] Joshua, S.C. and Garba, N.J. (1990). Estimating truck accident rate and involvement using poisson regression models. Transportation Planning and Technology 15, 41-58.
[5] Miaou S.P., Hu, P.S., Wright, T., Rathi, A.K., Davis, S.C., (1992). Relationship between track accidents and highway Geometric Design: A Poisson regression approach. Transportation Research Record 1376, 10-18.
[6] Lix, L.M., Ekuma, O., Brownell, M., and Roos, L.L (2004). A framework for modeling differences in regional mortality overtime. Journal of Epidemiology and Community Heath 58(5): 420-425.
[7] Yang,L., Lam, T.L., Lui, Y., Geng, K.W and Lui, C.D.(2005). Epidemiological profile of mortality due to injuries in three cities in Guangxi, China. Accident Analysis and Prevention 37 (1), 137-141.
[8] Tsauo, P., Elvik, R., Lopez, A. (1996). Age-period-cohort analysis of motor vehicle mortality in Taiwan, `774-1996. Accident Analysis and Prevention, 28 (5), 619-626.
[9] Kardara, C., and Kondakis, X. (1997). Road traffic accidents in Greece: Recent trends (1981-1991). European Journal of Epidemiology 13(7): 765-770.
[10] Lascala, M., Jadaan, K. and Homel, R. (1996). Epidemiology of Transportation Related Injuries in a Subsahara African.179-184.
[11] Evans, L. (2003). Estimating Accident Fatalities in Britain. Crowthonne, U.K: Transport Research Laboratory Report 445; 2003: 1-35.
[12] Atubi, A.O. (2012). Determinants of road traffic accident occurrences in Lagos state: some lessons for Nigeria. International Journal of Humanities and Social Science Vol.2, No.6, 252 - 259 (Special Issue).
[13] Oyedepo, O.J and Makinde, O.O. (2010). Accident prediction model for Akure-Ondo Carriageway, Ondo State Southwest Nigeria; Using Multiple linear regression. African Research Review Vol. 4(2): 30-49.
[14] Olushina, O.A and Mumimi, I.A.(2011). Multivariate regression techniques for analyzing auto-crash variables in Nigeria. Journal of Natural Sciences Research, 1(1):13-20.
[15] Aderamo, A.J. and Olatujoye, S. (2013). Trend in motocycle accidents in Lokoja. Nigeria. European International Journal of Science and Technology Vol. 2, No.6, 251-261.
[16] Nwobi-Okoye, C.C., (2011). Game Theoretic Analysis of Road Traffic Problems in Nigeria. African Research Review, 5(3):184-199.