

Factor Analysis of Road Accidents in FCT - Abuja, Nigeria

Yahaya Haruna¹, Dahida Terkuma².

¹ Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria

²Department of Mathematics, Faculty of Natural and Applied Science, Nigerian Army University Biu, Borno State, Nigeria

ABSTRACT: The study was designed to examine and analyze Road traffic crashes in FCT- Abuja, Nigeria. To construct groupings of road traffic crashes base on causes, form a correlation matrix to see if there is correlation among the various causes of road accidents and ascertain the effectiveness of the various measures taken to curb accidents in FCT, Abuja, Nigeria. We use biplot to form three groups from the causes of road traffic crashes which include: 1. Other causes, 2. Pedestrian crossing, 3. Reckless driving, Bad road, and inexperience. It was observed that the correlation between reckless driving, Bad road and inexperience was strong, positive and significant at 1% level which means that their variables can be used to predict (explain) one another. Also, it was observed that, the correlation between Pedestrian crossing, reckless driving, bad road and inexperience was weak, positive and significant at 5% level. And the correlation between reckless driving and other causes, Bad road and other causes, inexperience and other causes were insignificant, weak and positive. The KMO and Bartlett's test revealed that adequate sampling was used for this analysis. Principal component has suggested retaining two factors which accounted for 79.978% of the total variation.

KEYWORDS: Factors Analysis; Biplot, Pedestrian Crossing, Principal Component Analysis; Kaiser-Meyer-Olkin (KMO) and Bartlett's test.

I. INTRODUCTION

Road accident occur when a vehicle collides with another vehicle, pedestrian, animal, road debris or other stationary obstruction, such as tree or utility pole. Road accidents leads to death and disability as well as financial cost to both society and the individual involved. Road accidents in FCT have been a serious concern to most Nigerians in recent times as a result of it tremendous effect on human lives, properties and the environment as a whole. A lot of researchers have come out with the causes, effects and the recommendations to road crashes. These causes include pedestrian crossing, over speeding and inexperience etc¹. Despite all these efforts, every year, the Federal Road Safety Corps and other organizations would report that there is an increase in road accidents in the country¹.

Researchers in recent times have been modeling road accidents with crash prevention models in various parts of the world. However, it is extremely tedious to just apply model which have worked in other areas to data obtained from different countries due to the variations in the various factors pertaining in different countries². There has not been much statistical work conducted in road accidents in Nigeria and this might have been as a result of inadequate information available on road accidents and it impact on human lives and the environment as a whole in the country.

³has developed a forecasting model for traffic crashes for urban junctions with no road signs,⁴have also model traffic crashes and ⁵has modeled traffic crashes on rural highways. It is however surprising that in spite of the numerous factors identified by researches as the causes of road accidents and its consequences on human lives and properties; nobody has modeled the causes of road accidents and its contributions to the death and survival of casualties in Nigeria to authenticate the contributions of each of these factors to the casualties death. Furthermore, there has not been any work on the Likelihood of a casualty surviving in road accident or the possibility of a casualty surviving in a particular type of vehicle getting involved in a road accident. Road accident is defined as any activity which distracts the normal trajectory of a moving vehicle(s), in a manner that causes instability in the free flow of the vehicle. Vehicular accident has always been

attributed to human errors such as inexperience, over speeding, wrong overtaken among others. It has also been linked to poor road network, bad roads, cultism and the death-dying nature of some of the vehicles which ply the roads. There are many suggested solutions, various interventions by government, non – governmental organizations and other road stakeholders to curtail road accident and its effects on human lives and properties, it could be possible that these factors such as type and nature of the road, age and sex of casualty contribute casualty survival in road accidents in the country and have still not been considered. It is therefore in light of this that this research work is conducted in order to determine the appropriate ways to curb road accidents in FCT-Abuja, Nigeria. With the use of factor analysis to investigate the key causes of accidents that result in the death of people.

The research work seeks to carry out a factor analysis on the information on Road accidents recorded by the Federal Road Safety Corps Department of Corps Medical and Rescue (CMRS).

The Federal Road Safety Corps and the statisticians are needed to work hand in hand for accurate and adequate planning for decision making.

Road traffic accidents remain a dreadful plague all over the world. The trend has declined in most developed countries, with increase in developing countries. Road traffic accidents are endemic in Nigeria, with seasonal epidemics. According to¹, in the year 2000, a total of 80 road crashes were recorded, 240 people were involved, 139 were injured and 101 were killed in the year 2000. While in the year 2001, a total of 133 road crashes were recorded, a total of 561 people were involved, 312 people were injured, and 249 people were killed. Also recorded for 2002 shows a total of 525 people involved in 126 numbers of crashes, 184 people were killed, 341 injured. It was also observed that in 2015, 133 road crashes were recorded, 551 were involved, 360 were injured, while 191 were killed in FCT- Abuja, Nigeria.

Study has shown that 1 out of 3 and 4 out of 9 of the FCT-Abuja's population stand the risk of getting injured or killed on monthly basis from the road traffic accidents. Nigeria is ranked as one of the highly affected nations of the world. It is even more disturbing to note that the youth which are instrument and economic backbone of the nation, are mostly effected by this ravaging menace.

Hence, all hands must be on desk to see that the menace of road accident and their effects are deeply look into in other to proffer preventive measures and suggest possible strategies that will curb or reduce it to the barest minimum.

Therefore, it is important that a proactive action should be taken beyond the current measure by the government to check mate the spate of road crashes in FCT-Abuja. To be able to achieve the clear goal, the government needs to be graded with an in depth analysis of the situation obtaining with regard to the dynamics of road accidents in FCT-Abuja, Nigeria.

There are number of researchers who model road traffic crashes, such include⁶who used factor analysis to model road traffic crashes on example of elderly car drivers in Germany. Their research work reveals that the elderly have higher frequencies of certain violations of traffic rules than non-elderly car drivers. Also it shows that elderly car drivers are more often responsible for the accidents occurrence than non-elderly car drivers.

⁷on his research on factor analysis of road traffic accidents and casualty Risk in Scotland reveals that traffic accidents risk and traffic casualty risk are not associated with the same factors. But that places of resident is significant in determining casualty risk, but has no significant effect on accident risk.

⁸uses factor analysis to model road traffic crashes in Dehradum India. Their research work shows that accidents that happens in the night times on high ways are more critical than those that happens at other locations.

⁹In his research on road traffic crashes using factor analysis in north Carolina gives an indication that sleep related accident was the most severe accident among all other types of road accident

¹⁰showed through his research on road accidents using factor analysis shows that the driver and occupant of an older vehicle are usually at more risk of been killed than those in the new vehicle in case two vehicle collide.

¹¹in their research carried out in south Delhi, found out that among all the people who were killed in road accidents, 88.2 percent of them were males.

¹²proposed and sustained their researches that more males are killed in road accidents than females.

¹³In Wales and Great Britain assess the pattern of death of various age groups of casualties and their gender using factor analysis. This reveals that 40 percent of males and 30 percent of females who died in road accident were in age group 16 – 19 years.

¹⁴in their research as a contributory factor to death of casualties in road accidents is termed as “drink driving”. This could be due to inability of the drunk driver to control the vehicle as a result of sleeping.

¹⁴reveals in his research on factor analysis of road traffic crashes that 40 percent of those who died worn no seat belts and most of them were people within the age bracket 17 – 29 years. This gives the notion that no wearing of seat belt can also be a contributor to death of casualties in road accidents.

¹⁵Indicates that one of the most commonest thing identified as the cause of death in road accident is anoxia loss of oxygen supply. This cause a blockage in airways of casualties and the casualty immediately dies if not quickly attended to.

¹⁶in his research on social violence shows that, road accident fatality rate could be related to other aspects of social behavior such as murder and suicide rates in a given country.

¹⁷of Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, ROC made use of factor analysis to examine 28 groundwater samples collected from wells in the coastal Blackfoot disease area of Yun-Lin, Taiwan. Correlations among 13 hydro chemical parameters are statistically examined. A two-factor model is suggested and explains over 77.8% of the total groundwater quality variation. Factor I (seawater salinization) includes concentrations of EC, TDS, Cl_y, SO_{2y}, Na_q, K_q and Mg_{2q}, and Factor 2 (arsenic pollutant) includes concentrations of Alk, TOC and arsenic. Maps are drawn to show the geographical distribution of the factors. These maps delineate high salinity, and arsenic concentrations. The geographical distribution of the factor scores at individual wells does not reveal the sources of the constituents, which are instead, deduced from geological and hydrological evidence. The areas of high seawater salinization and arsenic pollution correspond well to the groundwater over pumping area. Over -pumping of the local groundwater causes land subsidence and gradual salinization by seawater. The over-pumping also introduces excess dissolved oxygen that oxidizes the immobile minerals, releases arsenic by reductive dissolution of arsenic-rich iron ox hydroxides and increases the arsenic concentration in water. The over extraction of groundwater is the major cause of groundwater salinization and arsenic pollution in the coastal area of Yun-L-in, Taiwan.

¹⁸: Celebrating the 100th anniversary of factor analysis, his research answer questions to why this approach has become in our days a principal statistical method of investigation in life sciences. What are the reasons of its wide dissemination in almost all scientific fields? Why so much efforts are directed towards new modifications and development of factor analysis methods? These questions can be answered from the positions of contemporary scientific methodology. While for most exact sciences, the using of differential equations, algebra, set theory, mathematical logic, and operational research is typical and usually sufficient, in such sciences as biology, psychology, sociology — methods intended to the analysis of multiple processes distribution and based on the probabilistic, rather than functional, homomorphism of the model and the object, are used. Therefore, the primary role passes on to the methods of mathematical statistics, theory of information, theory of random processes, etc. However, most of these methods have one crucial defect: they are not integral, giving answer to the question about the specifics and the reasons of organization of analyzed elements to a one system. The only acceptable method realizing the principles of system approach and meeting all above-mentioned requests is factor analysis as an integral statistical method, with its opportunity to define and evaluate the structural-functional organization of the system.

¹⁹in a research titled “an extended application of Factor Analysis in establishing peer groups among banks in Armenia “provided an historical backdrop to a modern-day analysis of banks in Armenia. Illustrated then are results from an extended application of Factor Analysis performed on a set of 17 banks and 13 financial variables, which are drawn from Armenia’s Arka News Agency data. Factor Analysis procedures was used to determine how banks can be classified based upon their disparate and common characteristics. Plots of outputs depict clusters of banks, which correspond with the statistical classification of banks based upon the 13 financial variables. Banks in any given cluster are deemed members of a ‘peer’ group which is a useful construct in performing comparative analyses among banks or for regulatory and policy-making purposes.



¹⁹performed Factor analyses on like banking data, it isolates and extends the Factor Analysis procedure to demonstrate and reinforce the efficacy of statistical procedures in the classification of entities. Specifically defined in this paper are the clusters of banks in Armenia which can be viewed as peer groups, polarized based on their financial characteristics embodied in the 13 banking variables for study from the Arka News Agency (Arka, 2001).

II. MATERIAL AND METHODS

This research work was carried out on road accidents in Federal Capital Territory (FCT), Abuja, Nigeria, from 2000 to 2015. The statistical Technique employed to analyze the data collected from Federal Road Safety Corp is factor analysis. The data used for this research work is collected base on causes.

Study location:The study was carried out in the Department of Statistics, Faculty of Science, University of Abuja, Nigeria.

Study Duration: The data was collected on a monthly basis from January, 2000 to December, 2015.

Sample

The collected data for this work is limited to sixteen years, i.e. from 2000 –2015and is restricted to road traffic crashes in the Federal Capital Territory (FCT) Abuja, Nigeria. The data was collected from Corps Medical and Rescue (CMRS) department of the Federal Road Safety Corps (FRSC) headquarters Abuja, Nigeria. The data iscollected base on causes of road traffic crashes, which include pedestrian crossing, reckless driving, inexperience, bad road and other causes.

Statistical analysis

Statistical package for social science (SPSS), was used to analyze the data. Correlation matrix was used to ascertain the extent of the relationship among the various causes of road traffic crashes.TheKaiser-Meyer-Olkin (KMO) and Bartlett's test was used to know whether adequate sampling was used for this analysis. Principal component was also employed to know the number of components to retain of the total variation.Thehypothesis (Ho) that the correlation matrix is an identity matrix was tested at 5% level of significance (Bartlett's test of Sphericity; $\chi^2 = 28.475$, p-value = 0.002), to know whether the correlation in the dataset are appropriate for factor analysis.

FACTOR ANALYSIS

Factor Analysis is a method of modeling observed variables and there covariance structure in terms of smaller number of underlying unobserved factors. It can be conceived as method for examining a matrix of correlation in search of clusters of highly correlated variables. Factor analysis has two main applications i.e reduce the number of variables and detect structure in the relationship between variables that is said to be classified.

The Factor Analysis Model

Consider the case of survey data, where we denote respondent i 's answer to survey question j , as X_{ij} ($i = 1, \dots, n; j = 1, \dots, k$). Factor analysis posits that X is a combination of unobserved factors, each written using:
$$X_{ij} = \lambda_{j1}\xi_{i1} + \lambda_{j2} + \dots + \lambda_{jp}\xi_{ip} + \delta_{ij} \quad \dots (1)$$

Where the λ terms are factor loadings to be estimated, and δ_{ij} is the measurement error in X_{ij} or that part of X that cannot be accounted for by the p underlying factors. It is possible to consider non-linear or multiplicative factor models, but the simple linear, additive structure in equation (1) is by far the more widely used factor analysis model.



Factor Loadings

The factor loadings λ are parameters to be estimated that tap how the unobserved factors account for the observed variables: the larger the values of λ , the more a particular variable is said to “load” on the corresponding factor. Noticing that the factor loadings k vary across survey items, but not across individuals. Put differently, items vary in the way they are explained by the underlying factors, but the relationships between underlying factors and observed responses is constant across individuals (hence the absence of an i subscript indexing k). Note also that there are fewer underlying factors than there are variables ($p < k$), consistent with the notion that like any statistical procedure, factor analysis is a device for “data reduction”, taking a possibly rich though unwieldy set of survey.

Estimation via Principal Component Analysis

In matrix notation, X_{ij} in equation (1) can be written as;

$$\Sigma = \Lambda\Lambda^1 + \psi \dots (2)$$

A popular and computationally cheap method for estimating the model parameters is via principal components analysis.

Principal component analysis is a method that transforms a set of highly correlated variables into a new set which are uncorrelated but contains almost all the information in the original data. The new set of variables are often fewer in number than the original variables.

The features of principal component analysis.

1. It incorporate no error term in its structure and hence it is a mathematical procedure.
2. It transforms a set of variables into a new set of uncorrelated variables.
3. It is appropriate when variables are unequally feting.

Measurement Errors

The terms for measurement errors simply reflect the idea that survey responses are not deterministically generated by the underlying factor structure. Like any statistical model, the factor structure is an approximation or a simplification that only captures so much of the survey responses under study.

III. RESULT

Table No. 1: Correlation Matrix

Correlation Matrix		RECKLESS DRIVING	PEDESTRIAN CROSSING	BAD ROAD	INEXPERIENCE	OTHER CAUSES
Correlation	RECKLESS – DRIVING	1.000	.496*	.689**	.617**	.346
	PEDESTRIAN CROSSING	.496*	1.000	.479*	.486*	.511
	BAD ROAD	.689**	.479*	1.000	.721**	.072
	INEXPERIENCE	.617**	.486*	.721**	1.000	.189



	OTHER CAUSES	.346	.511	.072	.189	1.000
Sig. (1-tailed)	RECKLESS – DRIVING		.025	.002	.005	.094
	PEDESTRIAN CROSSING	.025		.030	.028	.021
	BAD ROAD	.002	.030		.001	.396
	INEXPERIENCE	.005	.028	.001		.242
	OTHER CAUSES	.094	.021	.396	.242	

*Correlation is significant at the 0.05 level (1-tailed)

**Correlation is significant at the 0.01 level (1-tailed).

The correlation between Reckless Driving and Pedestrian crossing is weak, positive and significant at 5% level (.025).

The correlation between Reckless Driving and Bad road is strong, positive and significant at 1% level (.002).

The correlation between Reckless Driving and inexperience is strong, positive and significant at 1% level (.005).

The correlation between Pedestrian crossing and Bad road is weak, positive and significant at 5% level (.030)

The correlation between Pedestrian crossing and inexperience is weak, positive and significant at 5% level (.028)

The correlation between Bad road and inexperience is strong, positive and significant at 1% level (.001)

The correlation between other causes and Reckless Driving, Pedestrian crossing, Bad road and inexperience were insignificant.

Or

The correlation matrix in Table no 1 presents different levels of correlation between causes of road traffic crashes. There is strong positive relationship between reckless driving, Bad road, and inexperience. The relationships were also significant at 1% level, which means that their variables can be used to predict (explain) one another. Similarly, it was observed that the correlations between Pedestrian crossing, reckless driving, Bad road and inexperience were weak and significant at 5% level. It is also observed that the correlation between other causes, Reckless driving, Pedestrian crossing, Bad road, and inexperience, weak and insignificant.

Table No 2: Kaiser-Meyer-Olkin (KMO) and Bartlett’s test.

KMO and Bartlett’s Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.710	
Bartlett’s Test of Sphericity	Approx. Chi-Square	28.475
	Df	10
	p.value	.002



Hypothesis

Ho: $|p| = 1$

Vs

H1: $|p| \neq 1$

The null hypothesis(Ho) that the correlation matrix is an identity matrix was rejected at 5% level of significance (Bartlett's test of Sphericity; $\chi^2 = 28.475$, p-value = 0.002), this implies that the correlation in the dataset are appropriate for factor analysis. Also, "Kaiser-Meyer-Olkin statistic = 0.710" revealed that adequate sampling is being used for this analysis.

Table No 3: Communalities Initial extraction

Communalities		
	Initial	Extraction
RECKLESS DRIVING	1.000	.732
PEDESTRIAN CROSSING	1.000	.722
BAD ROAD	1.000	.869
INEXPERIENCE	1.000	.775
OTHER CAUSES	1.000	.901
Extraction Method: Principal Component Analysis.		

From Table no 3, we see that Bad road and other causes (.869 and .901) respectively, were best represented in the common factor space, this was because a high proportion of their variances was explained by the principal components.

FigureNo 1: Scree Plot.

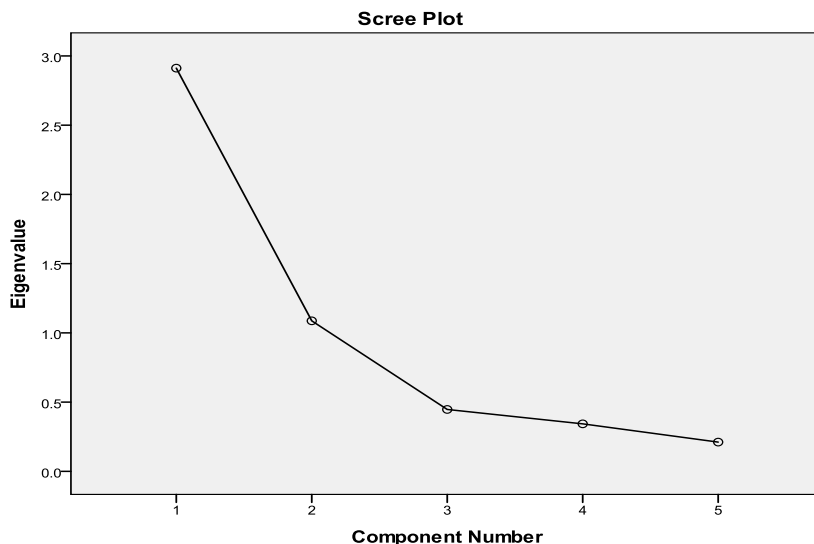




Table No 4: Eigen Value

Total Variance Explained			
Component	Initial Eigenvalues		
	Eigenvalue	Proportion	Cumulative
RECKLESS DRIVING	2.912	58.236	58.236
PEDESTRIAN CROSSING	1.087	21.740	79.976
BAD ROAD	.447	8.939	88.916
INEXPERIEN	.343	6.860	95.776
OTHER CAUSES	.211	4.224	100.000

Extraction Method: Principal Component Analysis.

The Eigen values, proportion and the cumulative are displayed in Table No 4, Considering the Eigen value-one criterion and the Scree plot in figure No 1, it would be reasonable to retain the first two PCs i.e. reckless driving and Pedestrian crossing. A commonly accepted rule says that it suffices to keep only PCs with Eigen values larger than 1, so the first 2 PCs can be retain to explain up to 79.976 percent of the total variability.

Table No 5: Eigenvectors: component factors estimates.

Component Matrix		
	Component	
	1	2
RECKLESS – DRIVING	.851	-.086
BAD ROAD	.830	-.425
INEXPERIENCE	.830	-.295
PED – CRS	.766	.366
OTHER CAUSES	.472	.824

Extraction Method: Principal Component Analysis.

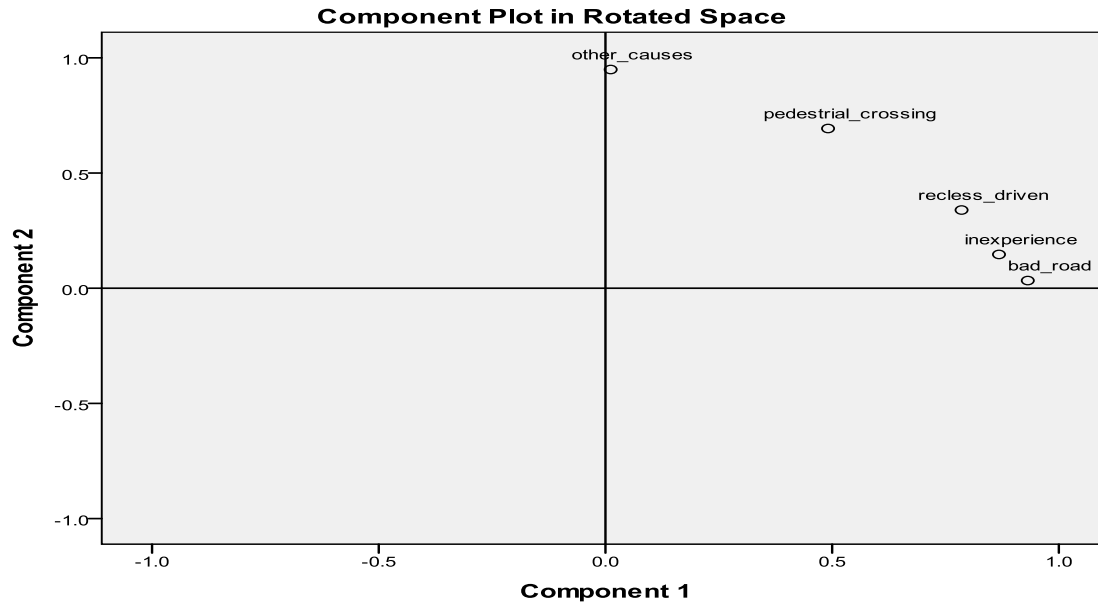
a. 2 components extracted.

Table No 6 above; concentrated on the two PCs that explains 79.976 per cent of the total variability of the data set.

The result shows that positive relationship exist among all the factors recorded under the first component, but majorly it identify, Reckless driving, bad road, and inexperience.

It also shows that negative relationship exist among reckless driving, Bad road and inexperience, But positive relationship exist among the remaining two factors under the second component.

Figure No 2: Component plot.



FigureNo 2 Above, Shows The Score Plot Which Was Used To Classify The Factors Into Three Groups, Which Are: 1. Other Causes, 2. Pedestrian Crossing And 3. Reckless Driving, Bad Road and Inexperience.

IV. DISCUSSION

Base on the analysis in this research work, three groups were formed from the causes of road traffic crashes which include: 1.Other causes, 2.Pedestrian crossing and 3. Reckless driving, Bad road and Inexperience. It was observed that a strong and positive correlation exist between Reckless driving, Bad road and inexperience. Also, it was observed that, the correlation between Pedestrian crossing, reckless driving, bad road and inexperience was weak, positive and significant at 5% level. And the correlation between reckless driving and other causes, Bad road and other causes, inexperience and other causes were insignificant, weak and positive. The KMO and Bartlett’s test revealed that adequate sampling was used for this analysis. Principal component has suggested retaining two factors which accounted for 79.978% of the total variation.

V. CONCLUSION

Based on the data analysis in this research work, and the objectives, the following conclusions were made:

1. To construct groupings of factor analysis of road traffic crashes base on causes. Three groups were formed which include: 1. other causes, 2. Pedestrian crossing and 3. Reckless driving, Bad road and Inexperience.
2. To form a correlation matrix to see if there is correlation among the various causes of road traffic crashes. There is correlation among the various causes of road traffic crashes.
3. To ascertain the effectiveness of the various measures adopted to curb auto crashes in FCT-Abuja, Nigeria. Despite all the measures adopted to curb road traffic crashes, it was observe that road traffic crashes keep increasing on daily basis.
4. To suggest possible ways to curb road traffic crashes in FCT Abuja. We suggest that there should be provision of good roads and regular maintenance. Also, the law enforcement agencies should be very vigilant on drivers who violate traffic rules.

REFERENCES

- [1]. Federal Road Safety Corps FCT, Abuja, Nigeria (2009), Annual Report, FCT Abuja.
- [2]. Fletcher et al (2006), Road Accident with crash prevention model.
- [3]. Salifu (2004), Forecasting Model for Traffic Crashes for Urban Junctions with no road signs.
- [4]. Afukaar and Debrah (2007) Road traffic Accidents.
- [5]. Ackaah (2011) Road traffic crashes on rural highways.
- [6]. Otte et.al (2008) factor analysis of road traffic crashes on example of elderly car drivers.
- [7]. David I. W (2002), Factor Analysis of Road Traffic Crashes and casualty risk in .Scotland.
- [8]. Sachim et al. (2009 – 2011), Factor Analysis of Road Accident in Dehrandum India.
- [9]. Allen et.el (1995) Road traffic crashes using factor analysis, North Carolina.
- [10]. Broughton, (2007), Factor analysis of road traffic crashes.
- [11]. Kumar et al. (2008), Road Accident in South Delhi.
- [12]. Salgado & Columbaje (1998), Road Traffic Crashes based on Gender.
- [13]. Department of Transport (2006), Pattern of Death of all age groups, Wales and Great Britain.
- [14]. Clerk et.al (2007) contributory factors to death of casualties in road accident.
- [15]. British Red Cross (1997), Causes of road accident.
- [16]. Witlock (1971), Social Violence.
- [17]. Chen-Wuing L., Kao-Hung L., Yi-Ming K. (2003). Application of factor analysis in the assessment of ground water quality in a Blackfoot disease area in Taiwan. Department of Bioenvironmental System Engineering, National Taiwan University, Taipei 10617, Taiwan ROCProc. Estonian Acad.Sci. Chem. 2001.
- [18]. Kaplunovsky, A.S. (2004), Why using factor analysis? Research center of Quantum Communication Engineering Horon Academic Institution, London.
- [19]. Safdari (2003, 10). An extended Application of Factor Analysis in establishing peer groups among banks in America. By Cyrus publication: Journal of Academic, of Business and Academic.