KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF SCIENCE

FACULTY OF DISTANCE LEARNING

DEPARTMENT OF MATHEMATICS

TOPIC

IDENTIFICATION OF FACTORS THAT CAUSE SEVERITY OF ROAD ACCIDENTS IN GHANA. A CASE STUDY OF THE NORTHERN REGION.

NAME: ABDUL-RAHAMAN HAADI

2012

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY
COLLEGE OF SCIENCE,

FACULTY OF DISTANCE LEARNING,

DEPARTMENT OF MATHEMATICS.

IDENTIFICATION OF FACTORS THAT CAUSE SEVERITY OF ROAD ACCIDENTS IN GHANA. A CASE STUDY OF THE NORTHERN REGION.

A DISSERTATION SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES (KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY) IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE MASTER OF SCIENCE DEGREE (MSC) IN INDUSTRIAL MATHEMATICS

ABDUL-RAHAMAN HAADI

2012

DECLARATION

This thesis is submitted to KNUST, School of Graduate Studies through the College of Physical Science. I hereby declare that this thesis is my independent work and has not been accepted in any previous application for a degree here or elsewhere. This thesis presents results of original research undertaken by the undersigned. Information taken from other works has been specially and duly acknowledged.

………………………. ……………..

ABDUL-RAHAMAN HAADI DATE……………………..

(STUDENT)

……………………………………….

MR. ISAAC AKPOR ADJEI DATE……………………..

(SUPERVISOR)

………………………………………

MR. KWEKU DARKWAH DATE………………………

(HEAD OF DEPARTEMENT)

ACKNOWLEDGEMENT

All thanks to the Almighty for seeing me through all this while for without His support I wouldn’t be where I am.

Special thanks to my Supervisor Mr. Isaac Adjei for his exhaustive checks and proof-readings of my scripts.

My heartfelt gratitude goes to my family; my Mum and Dad, and my Siblings for the support they gave me throughout this program. I can’t forget to mention my Sweetheart Alhassan Mubarika whose support has been enormous may the Almighty Allah richly bless her.

My deepest appreciation goes to Joyce Kande for the patience, support, understanding and endless encouragement she gave me all that while.

Big thanks to Mr. Christopher Kansanba (Pro-Credit), Mr. Raymond Kansanba (Garden City University Kumasi) and his wife Rowena, Mr. Abdul-Rahaman Abdul-Aziz (K-Poly), Mr. Issahaku Baba Adam (Tamale Teaching Hospital), Mr. Timothy Abeo, Mr. Mohammed Awal Kpahambang, Mr. Nabia Alhassan John, Mr. Abdul-Aziz Ibn Musah, Mr. Abraham Alidu, Mr. Tuahiru Maahi, all of Tamale Poly, Mr. Alhassan Abdul Barik (UDS Navrongo Campus), my one and only Imam Alhassan Faisal (UDS Navrongo Campus) and everybody who in diverse ways helped in the success of my program and this work. MAY THE ALMIGHTY ALLAH RICHLY BLESS YOU ALL.

Finally, to my two great friends Abdulai Abdul-Rafik (GHS) and Ibn Tuahir Kassi Abdul-Latif (SOS Village Tamale), guys keep it tight.

DEDICATION

This work is dedicated to the Almighty Allah, my entire family and friends for the enormous support, love and encouragement they gave me in the course of writing the thesis.

ABSTRACT

The problem of deaths and injury as a result of road accidents is now acknowledged to be a global phenomenon. As a result authorities in virtually all countries of the world are now concerned about the growth in the number of people killed and seriously injured on their roads including Ghana.

The study objective was to identify the variables that mainly contribute to accident severity in the Northern Region and to describe the impact of these variables.

This study applied logistic regression to accident-related data collected from MTTU Northern Region traffic-police records in order to examine the contributing factors to accident severity. A total of 398 accident data from 2007-2009 was used.

The accident severity (dependent variable) in this study is a dichotomous variable with two categories, Fatal or Non-fatal. Among the variables obtained from police-accident reports, two independent variables were found most significantly associated with accident severity; namely, Over loading and Obstruction.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

WHO	World Health Organization
UNICEF	United Nations Children's Fund
UN	United Nations
MTTU	Motor Traffic and Transport Unit
GDP	Gross Domestic Product
NRSC	National Road safety Commission
SPSS	Scientific Product and Service Solution
BRRI	Building and Road Research Institute
DVLA	Driver and Vehicle Licensing Authority
GHA	Ghana Highways Authority
DUR	Department of Urban Roads
DFR	Department of Feeder Roads
GNP	Gross National Product
VMT	Vehicle Miles of Travel
HALs	High Accident Locations
RTAs	Road Traffic Accidents

CART	Classification and Adaptive Regression Trees
ARM	Accident Risk Model

CHAPTER 1

1.0 INTRODUCTION

It is well known that traffic accidents are of high importance to the public health spectrum in the world. Moreover, in developing countries such as Ghana, the mortality rates from road traffic accidents are rather high compared to other countries in this region. Not only the majority of the people killed and seriously injured significantly affect the quality of life of the citizens but it also inhibits the economic and social development of this country.

In order to use accident data to establish the "prevention measures" to efficiently support the decision of policy makers, an effective "accident data gathering system" is the most important pre-requisite in solving the accident problems. Since there are many factors responsible for the high accident rates, therefore process of accident data collection should cover the important accident information such as road type, environmental features and location, as well as the details of the vehicle/driver and casualty information. Moreover, incomplete accident data gathered by responsible agency, added to the increasing of difficulties in analyzing motorcycle accidents.

This study intends to identify factors that cause severity of accidents in the Northern Region and will use logistic regression model to help achieve that.

1.1 BACKGROUND

The problem of deaths and injury as a result of road accidents is now acknowledged to be a global phenomenon. As a result authorities in virtually all countries of the world are now concerned about the growth in the number of people killed and seriously injured on their roads.

Road traffic injuries are a major but neglected public health challenge that requires concerted efforts for effective and sustainable prevention. Of all the systems with which people have to deal every day, road traffic systems are the most complex and the most dangerous. Worldwide, an estimated 1.2 million people are killed in road accidents each year and as many as 50 million are injured. (WHO: World report on road traffic injury prevention, 2011). Projections indicate that these figures will increase by about 65% over the next 20 years unless there is new commitment to prevention. Nevertheless, the tragedy behind these figures attracts less mass media attention than other, less frequent types of tragedy. (WHO: World report on road traffic injury prevention, 2011)

The Global status report on road safety is the first broad assessment of the road safety situation in 178 countries, using data drawn from a standardized survey. The results show that road traffic injuries remain an important public health problem, particularly for low-income and middle-income countries. (WHO: Global status report on road safety, 2009). Pedestrians, cyclists and motorcyclists make up almost half of those killed on the roads, highlighting the need for these road users to be given more attention in road safety programmes. The results suggest that in many countries road safety laws need to be made more comprehensive while enforcement should be strengthened. The Global status report on road safety results clearly show that significantly more action is needed to make the world's roads safer. (WHO: Global status report on road safety, 2009)

Accident severity is of special concern in traffic safety, as many efforts address accidents tend to be measured not only to prevent accidents but also to reduce the severity of accident. One way to do so is to identify the most probable contributing factors that affect accident severity.

Media reports reveal that Ghana's road accident's is very high, compared to other developing countries. In 2001, Ghana was rated as the second highest road traffic accident-prone nation among six West African countries, with 73 deaths per 1000 accidents. (Sarpong, S. (2011). Road accident on the increase. Retrieved December 21, 2011.)

Road accidents have great effect on any country's economy especially, a developing country, like Ghana. Statistics from the National Road Safety Commission shows that, Ghana loses about 1.7 per cent of its Gross Domestic Product which is over 230 million dollars every year. Beside the loss of lives, $165 of the Country's Gross Domestic Product (GDP) to road accidents. (Ghana news agency. (2010). Road accidents cost Ghana millions of dollars. Retrieved November 11, 2011, from Ghanaweb web site)

Road casualties represent a cost to the country, whether dead or alive since victims become physically disabled and therefore can't work effectively, to increase productivity and development in the country.
Many people continue to be negligent and ignore the dangers involved in driving and few precautions are observed by motorists.

The most talked about and serious cause of accidents nowadays is the use of cell phone while driving. Drivers using cell phones, tend to miss the traffic signals since there is the absence of concentration on the road. The driver may be looking on the road but can easily get distracted by the process of answering or dialing a number and this can result in fatal accidents on the road.

Another cause of road accidents is over speeding. Motorists who go beyond or below the prescribed speed limits on road could hit other vehicles and they could be hit as well.

Statistics show that about 30% of accidents on the roads are caused because of speeding. (Akongbota J. (2011). Reducing accidents on our roads. Retrieved December 21, 2011, from government of Ghana official portal web site.). Speed limit also applies to conditional weather. Drivers have to slow down when it is raining or foggy, if not once they speed up; they face great responsibility of losing control over the vehicle.

Also, driving while intoxicated can cause fatal road accidents. Alcohol and drugs can affect ones driving skills. It causes general impairment of the brain and the function of nervous system.

Drivers who drive tired or exhausted are mostly sleepy, agitated and aggressive on the roads. Fatigue can affect their clear thinking if they encounter a glitch on the roads and this can lead to accident.

Some drivers disregard road rules since they are more concerned of getting to their destination than of how they would get there. Such drivers would violate road rules and laws just to get to where they are going. All these, would definitely cause serious road accidents.

As much as accidents are caused by various factors, there are always preventive measures or steps that can be taken to prevent them.

This study aims at examining factors, that have a higher potential for serious injury or death, such as accident location, type, and time; collision type; age and, nationality of driver at fault, and his licensing status; and vehicle type. The Logistic regression is used in this study to estimate the effect of the statistically significant factors on severity. Logistic regression and other related-categorical-data regression have often been used to assess risk factors for various diseases. However, it has also been used in transportation studies.

1.2 STATEMENT OF THE PROBLEM/JUSTIFICATION

A road traffic accident is defined as any vehicle accident occurring on a public highway. It includes collisions between vehicles and animals, vehicles and pedestrians, or vehicles and fixed obstacles. Single vehicle accidents, which involve a single vehicle, that means without other road user, are also included (Safe car guide, 2004).

At all levels, whether at national or international level, road traffic accidents continue to be a growing problem. In connection with this, according to a World Health Organization/World Bank Report, deaths from non-communicable diseases are expected to grow from 28.1 million a year in 1990 to 49.7 million by 2020, which is an increase in absolute number of 77%. Traffic accidents are the main cause of this rise. Road traffic injures are expected to take higher place in the rank order of disease burden in the near future.

The level of vehicular accidents in Ghana is alarming and particularly sad since they are all preventable. Of all the systems that the people of Ghana have to deal with on a daily basis, road transport is the most complex and the most dangerous. Every day people are killed and injured on our roads. (Kwame Y. (2011, August 21). Congratulations GPRTU, but Muslim drivers also drink. Retrieved December 21, 2011.)

According to the World Health Organization, African nations have the world's highest road traffic injury mortality rates. Whiles the average death rate from road accidents in developed nations is around 20.8 per 100,000 populations that of Ghana is 29.9 per 100,000.

From January to March this year, Accra alone recorded 1,417 motor accidents involving 2,125 vehicles. According to the Motor Traffic and Transport Unit (MTTU) of the police during this period, there were 78 fatalities, 373 serious injury cases and 966 minor cases in which vehicles ran into other vehicles. (Sarpong, S. (2011). Road accident on the increase. Retrieved December 21, 2011.)

The National Road Safety Commission is aiming at reducing accident fatality rate to a single digit by 2015 from its present rate of 19 fatalities per 10,000 vehicles.

This study concentrates on the number of injured people per accident and will attempt to address the following questions:

- What are the variables that mainly determine the severity of road accidents in Ghana?
- What is the impact of these variables on severity of road accidents?

1.3 OBJECTIVES OF THE STUDY

The general objective of this study is to identify factors that cause severity of road accidents in the northern region.

The specific objectives are:

- To identify and describe the major variables (factors) that contributes to severity of road accidents.
- To analyze the impact of these variables (factors) on severity of road accidents.

1.4 THE STUDY AREA

The Northern Region has a total land area of about 70,384 sq. km which is 29% of the land area of Ghana. It is located on 0 longitude (Greenwich Meridian) and between latitude 8 30" and 10 30" N and lies completely in the savannah belt. It has Togo and La Cote D'Ivoire to the East and West respectively, as its international neighbours. Further south, the region shares boundaries with Brong Ahafo and the Volta Regions, and to the north it shares borders with the two Upper Regions. (Ghana Health Service website, retrieved on November 22, 2011.)

The 2000 census projection report gives the regional population as 1,820,000 at growth rate of 2.9% giving a population of 2,090,399 in 2005. It is divided into eighteen political/administrative districts headed by the District Chief Executives. The districts are further subdivided into seventy five health sub-districts. Most of these correspond with the local council zones.

This is tropical with temperatures ranging from a low of 14 degrees Celsius night temperature during the hamattan season to as high as 38 degrees Celsius during the hot dry season. The rains begin lightly in April and rise steadily to peak in August-September and gradually decline by the end of October. The dry hamattan winds engulf the whole region between December and February. In recent years the rains have been starting late, in May and peaking later in September-October.

The 2000 census report puts the population of the region at 1,820,806. At a growth rate of 2.9 per annum the estimated population for 2005 is 2,090,399. (Ghana Health Service website, retrieved on November 22, 2011.)

This population is characteristically distributed in small settlements with populations of 200 - 500 people. There are over 5,000 settlements in the Region, out of which 54.4% have population less than 200 people. The distances between settlements are far apart. This peculiar pattern of distribution of population in the Region has adverse implication for service delivery, as Service Delivery Health Teams (SDHTs) going on out-reach travel long distances only to reach a small proportion of their target population.

A reasonable proportion of the population is in “overseas” areas in seven of the eighteen districts namely, East Gonja, West Gonja, West Mamprusi, Nanumba South, Gushegu, Karaga and Tolon/Kumbungu districts. These populations can only be assessed from neighboring regions/districts or only during the dry season In the West Gonja and East Gonja districts, several villages are completely surrounded by the Volta Lake.

Poverty is high and widespread and many cannot afford the cost of basic health services Agriculture remains the predominant sector. With over 90% of the productive age group being peasant farmers. Mechanized agriculture is possible on this terrain although limited in practice because of the high cost of inputs. However, the peasant farmer produces the bulk of the cereals, tubers and groundnuts in the region. Sheanut is the most important cash crop in the region.

Cotton Ginnery is perhaps the only industrial sector with a high out-put level. Notwithstanding the low activity in this sector the establishment of the Intermediate Technology Transfer Unit has been a booster to entrepreneurs who depend on it for the manufacture of spares, tools etc. for their light industries. Leather tanning is also done on a large scale. A number of mining activities have sprung up in some districts, notably Bole but this is still on a low scale.

The state of the roads in the Region is generally bad. The only stretch roads that are tarred are the Buipe through Tamale to Bolgatanga stretch, a distance of about 270 Km and Tamale to Yendi stretch. Most roads are not motorable in the rainy season thus hampering outreach and other health programs. Bicycles and motorbikes are therefore more effective for out-reach activities in the Region.

There has been a significant improvement in telecommunications over the past three years; most of the district capitals can now be reached by telephone: eight out of the eighteen can be reached directly by phone. One out of the eighteen has not yet been connected to the national grid.

The region is among those with the lowest school enrolment rate, highest dropout rate and highest illiteracy rates in the country.

1.5 RESEARCH METHODOLOGY

The data for the analysis of the study were based on accident data from the MTTU Northern Regional office of Ghana Police Service. The data collected was on vehicular accidents that have occurred in the region for the period, 2007–2009, indicating their sex, age, age of vehicle, type of vehicle, Estimated Speed, Road surface condition when accident occurred, Weather condition when accident occurred, Traffic lighting condition when accident occurred, Drivers familiarity of route, Type of driving license, Scene of accident and Cause of accident.

The National Road Safety Commission (NRSC) was also contacted for some literature concerning accidents in Northern Region and Ghana as a whole. Some of the materials for the study were obtained from the Mathematics department library and also from the internet.

To achieve the objectives of this research, the logistic regression model would be employed as the main statistical methodology for analyzing the data. It will be used to predict the relative likelihood of the severity of the accidents. Logistic regression portrays the connection between a categorical response variable and a set of predictor/independent variables.

Also the Wald tests, together with the Deviance, and Chi-square test will be used as criteria to include or remove independent variables in the model. The dependent variable in this research is Accident and of dichotomous type and stands for accident severity. Each accident in the sampled data will be categorized as either injury or fatal and SPSS software package would be used in obtaining the analysis.

1.6 LIMITATIONS OF THE STUDY

This study is based on a secondary data obtained from Motor Traffic and Transport Union (MTTU). The data contains records of road traffic accidents and the number of injuries (which includes fatalities) in each accident.

Some data were collected by interviewing drivers and the accuracy of the data is questionable to some extent. There were also incomplete and inconsistent records.

Regardless of the above limitations, the resultant error is assumed to be insignificant due to the large sample size used. In addition, only those records found complete and consistent are considered for the study.

1.7 ORGANIZATION OF THESIS

Chapter 1 presents the background of the study, problem statement, objectives of the study, study area, limitations of the study and the organization of the thesis.

Chapter 2 describes recent trends in fatalities of road traffic accidents, factors influencing severity of road accidents, and literature on previous research studies attempting to identify these factors.

Chapter 3 discusses the sources of data, and statistical methodology used in this study.

Chapter 4 presents and discusses the results of the analysis.

Chapter 5 presents the conclusions and recommendations of the study.

CHAPTER 2

LITERATURE REVIEW

2.0 INTRODUCTION

The topic of crash severity has been of interest to traffic safety community because of the direct impact on occupants involved. The way forward would be to identify factors contributing to either a more or less severe crash. The approaches used to model injury severities vary from one to another, depending on the purpose of the study and data availability. This chapter provides a review of recent trends in motor vehicle crashes, factors influencing motorcycle traffic safety and various methodologies commonly used in traffic safety studies as well as a discussion of findings in the field of traffic safety. The literature review focuses primarily on road safety situations in Ghana, the types of analyses available and the factors that relates to severity of road accidents.

2.1 A REVIEW OF ROAD SAFETY SITUATIONS IN GHANA

2.1.1 ROADS SAFETY SITUATION

The number of motor vehicles in Ghana is increasing rapidly and, coupled with population growth, is contributing to a rise in the number of road traffic injuries and fatalities. Road safety has become a major national issue receiving front-page coverage in the press and National TV news on a regular basis. Fortunately, the government and donor community have reacted quickly and increased funding to the National Road Safety Commission (NRSC), enabling the NRSC to expand and implement new targeted road safety initiatives. The Danish International Development Assistance, Danida, has been a primary supporter of government road safety activities in Ghana.

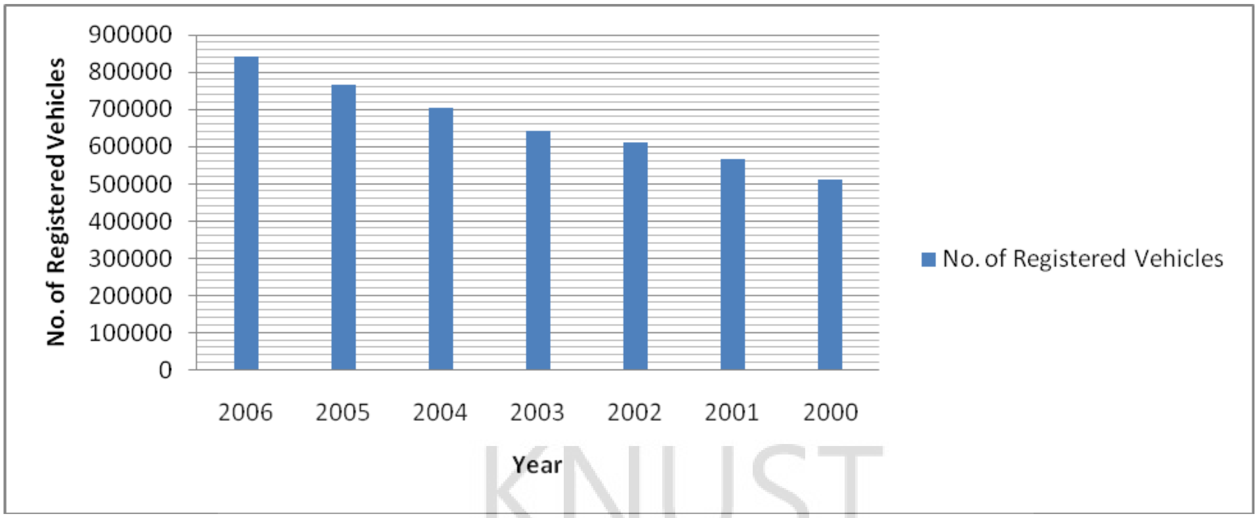

Figure 2.1 Number of Registered Vehicles (2000 – 2006)

Road crashes kill an average of four persons daily in Ghana. In 2005, the latest year for which statistics are available, the number of road crashes increased by 16% relative to 2004. (National Road Safety Commission website, www.nrsc.gov.gh/statistics, retrieved on November 7, 2011) The regions Ashanti, Eastern, Gt. Accra, Central and Brong Ahafo Regions account for more than 70% of the total number of crash fatalities.

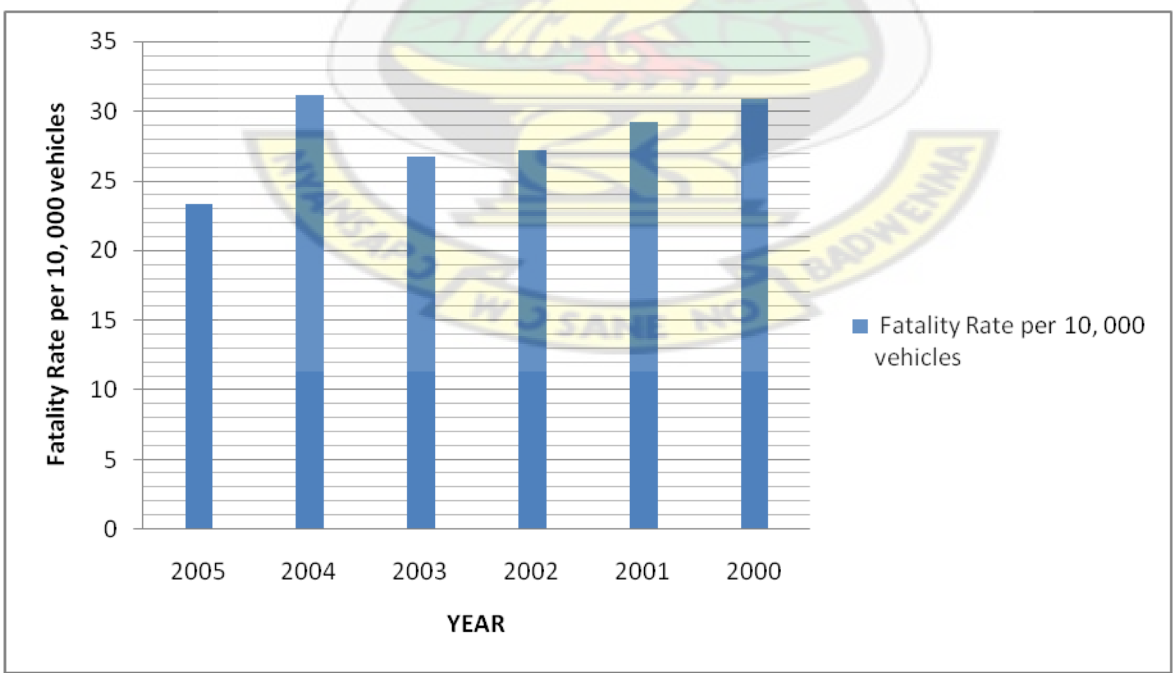

Figure 2.2 Road Traffic Crash Fatality Rate per 10,000 Vehicles (2000 – 2005)

Some 70% of crashes occur on flat and straight roads. Speeding is a major cause of crashes, accounting for over 50% of reported crashes. Buses and mini-buses cause 35% of fatal crashes while cars are responsible for 32%.

Road users between 16-45 years are the most vulnerable group and account for 58% of total road crash fatalities from 2002-2006. 70% of persons killed in road crashes are males. The age groups from 0-5, 46-65 and over 65 years also accounted for a 20.8%, 16.7% and 4.6% respectively of the total fatalities during the same period.

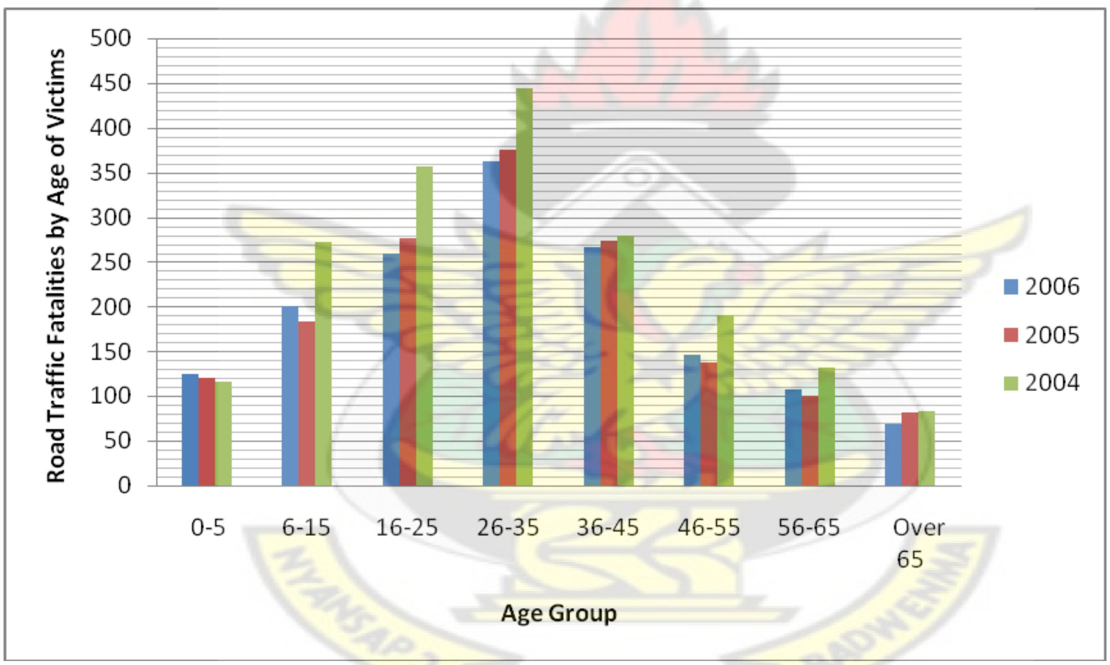

Figure 2.3 Road Traffic Fatalities by Victims Age (2004 – 2006)

Pedestrians continue to be the most vulnerable road user group, accounting for more than 40% of the annual road crash fatalities. 21% of pedestrians killed annually are children below the age of 16 years of age. This was followed by occupants of buses/Mini buses, cars and Heavy Goods Vehicles (HGV) also recording 23%, 12% and 10% respectively.

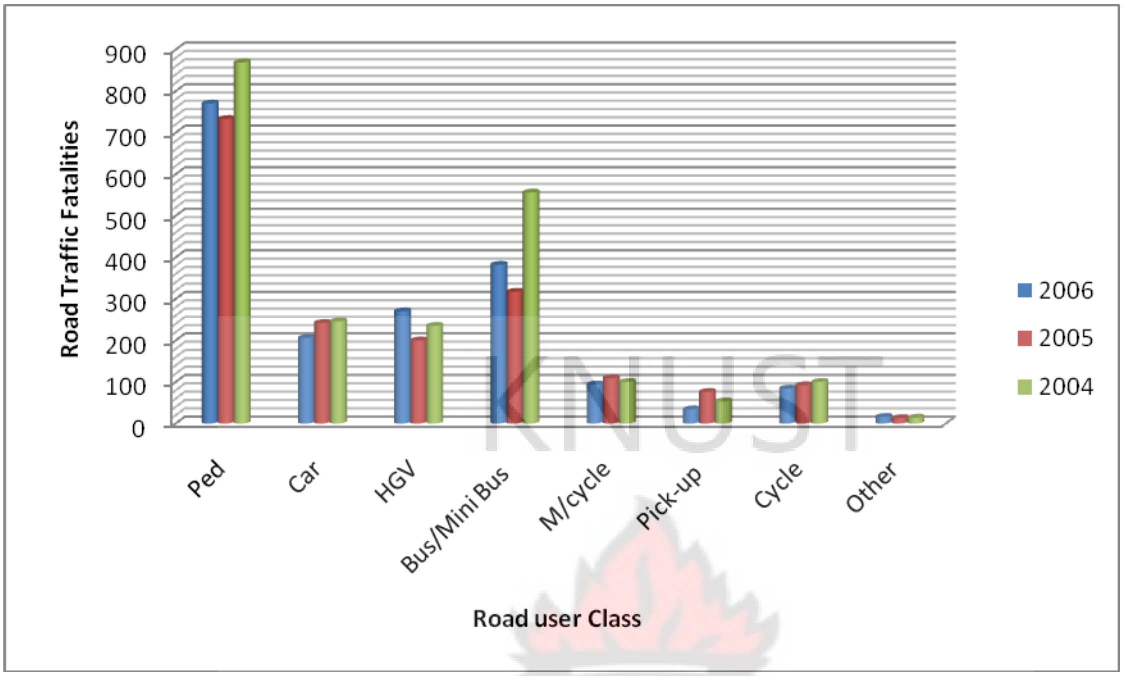

Figure 2.4 Road Traffic Fatalities by Road User Class (2004 – 2006)

Although fatality numbers are growing, crash risk has remained stable over the past six years. The statistics should be viewed with caution as the quality of the national data is affected by under-reporting (including both non-reporting, because not all crashes are reported, and under-recording arising from incomplete retrieval of reported crashes from police files). Whereas the actual degree of under-reporting is not currently known, the level of under-recording of the data has been reviewed and generally improved over the years.

Data management systems are improving as IT systems become more widely available and this will have an impact on the crash reduction programme from better informed safety programmes.

2.1.2 EFFECTS OF THE MENACE

Indeed this conquerable foe is gluttonously devouring our human and economic resources. Precious lives are lost thereby dwindling down our scarce labour force in the country, causing a stir in our Gross Domestic Product (GDP). Continual media reports reveal that Ghana's road accident is oddly high among developing countries. In 2001, for example, Ghana was rated as the second highest road traffic accident-prone nation among 6 West African countries, with 73 deaths per 1000 accidents.

Road accident increased in 2001 to 11,291 with 1660 fatalities, while decreasing slightly in 2002 to 10, 718 but with 1665 fatalities. Though road accidents further declined to 10,644 in 2003, fatalities piercingly rose to 11,718. Accra alone recorded, from January to March 2003, 1,417 motor accidents involving 2,125 vehicles. According to the officials of Motor Traffic and Transport Unit (MTTU) of the police, during this period, there were 78 fatalities, 373 serious injury cases and 966 minor accidents among others. The figure of road accidents in the country sharply went up to 12, 164 in 2004 but decreased to 11, 305 in 2005. Ghana records about 10,000 fatal road traffic accidents, every year, out of which 1,600 people perished while 15,000, are seriously injured, robbing the nation of some precious lives. Such persons may die or become incapacitated, denying them the ability to contribute to the nation's development meaningfully. Besides, Ghana loses an amount of GH¢165, 000, representing 1.6% of our GDP, yearly in solving road accident situations such as medical expenses of victims, damage to vehicles and insurance cost among others. Our dear nation in all loses about 2% of her GDP annually due to road accidents.

2.1.3 ROAD SAFETY COORDINATION AND STAKEHOLDERS

In the past, road safety activities were carried out by individual departments within the Ministry of Roads and Transport: Ghana Highways Authority, Department of Urban Roads, Department of Feeder Roads, Driver Vehicle Licensing Authority, together with the National Road Safety

Committee. In 2000, the Committee became a Commission by act of Parliament and the Commission coordinates safety activities within Ghana.

The Government of Ghana, acting through the Ministry of Transport and Communications, established the National Road Safety Commission (NRSC) to develop, promote and coordinate the National Road Safety Strategy.

The National Road Safety Strategy provided the NRSC with guidelines for its work in the 5-year period 2001-05.The purpose of the strategy was to break the upward trend in crashes, injuries and fatalities and create a basis for concrete, sustainable crash reduction by 2010. The overall target was a 5% reduction in road fatalities from 1998 as the base year to 2005 and a further 15% reduction before the end of 2010. A new 5-year strategy is currently being prepared.

The NRSC has identified seven major public road safety stakeholders who will be the mainstays for the implementation of the strategy. These are:

- The National Road Safety Commission and its Secretariat (NRSC);
- The Building and Road Research Institute (BRRI);
- The Driver and Vehicle Licensing Authority (DVLA);
- The Motor Transport and Traffic Unit - the traffic police (MTTU);
- The Ghana Highways Authority (GHA);
- The Department of Urban Roads (DUR);
- The Department of Feeder Roads (DFR);

Six major road safety problems have been selected as focus areas for the strategy:

- Accident black spots in urban areas and villages;

- Pedestrians;
- Children;
- Professional drivers;
- Speeding;
- Drunken driving;

The NRSC has also identified other stakeholders from the public and the private sectors who can contribute to improving road safety. Among these are the Transport Coordinating Council and the oil marketing companies. Private companies have in the past provided funding for road safety campaigns and road safety education programmes for school children. Shell and Exxon have successfully demonstrated that road safety training for drivers can contribute to reduced crash costs for the companies.

Ghana established a Road Fund in 1997. The monies for the Fund are derived from levies on fuel, tolls and vehicle licenses. They are managed by a Board consisting of 13 members, the majority of whom are from the private sector. The Fund can be used for road safety activities and there is tentative agreement to make 900 million (approximately US$ 150,000 at year 2000 exchange rates) available for road safety provide a reasonable action plan is submitted. Up-to-date road crash cost figures are not available for Ghana but using 1% of GNP as a typical minimum estimate gives a figure of $72.6 million per annum.

2.1.4 THE ROOT CAUSE OF THE MENACE

The most common known causes of road traffic accidents in Ghana include gross indiscipline on our roads, over-loading, fatigue driving, drunk driving and over-speeding. Statistics show that 60% of road accidents are caused by drunk driving and over-speeding. The latter alone constitutes about 50% of road accidents in the country. The poor nature of some of our roads, poor maintenance of vehicles, disregard for traffic regulations by most drivers and indiscriminate use of the road by some pedestrian are some of the other causes of motor accidents in the country.

2.2 FACTORS AFFECTING SEVERITY OF ACCIDENTS

Robert, B. (2000), analyzed the relationship between road infrastructure and safety by using a cross-sectional time-series data base collected for all 50 U.S. states over 14 years. Data on total fatalities and total injuries by state were collected. Data on road infrastructure included total lane miles (excluding local roads), average number of lanes by functional road category (interstates, arterials, and collectors), percent of center-line miles with a given lane width by road category, and the fractional percent of each road category in a given state (including local roads within the denominator). Interstates are controlled access highways built to the most rigorous and consistent design standards. Arterials are generally major multi-lane or intercity roads, perhaps with some controlled access, but generally not. These also tend to be major connector roads within cities and suburban areas. Collector roads are smaller scale roads that generally connect local distributor roads with arterials. A casual interpretation of the trends and those for total fatalities would suggest that as highway facilities are upgraded, there are reduced fatalities. In addition, estimates of seatbelt usage, by state, were used to control for the effects of increased seatbelt use. The analyses also attempts to control for seatbelt effects by including dummy variables for those states with either primary or secondary seatbelt laws. Data on total population, vehicle miles of travel (VMT), per capita income, alcohol consumption and population by age cohorts was also collected. These are used in the models primarily to control for other factors that are likely to affect fatalities and injuries. The occurrence of traffic crashes and the resulting injuries and fatalities are Poisson distributed. The use of a Poisson regression is usually affected by over-dispersion in the error term due to the inequality of the mean and variance within the data. This is easily corrected by using a negative binomial regression. A number of different models were estimated using the data described above. The key variables of interest are the infrastructure variables. Other variables known to affect crashes are also included, specifically age cohorts, per capita income, state population, and VMT. VMT and population size could not be included in the same model due to high collinearity between them. Separate models for each were therefore estimated. The dependent variables were the total deaths (DEATHS) and total injuries (INJURED) from traffic-related crashes. Models containing all the relevant infrastructure variables and models without the lane width variables were developed. According to the results of his study, total lane miles are found to be highly significant across some of models for both total fatalities and injuries. No significant effect is found for increases in the average amount of interstate lanes on fatalities. The percent of lane miles by each road category shows that those states with more lane miles of interstate (relative to other categories) have a statistically significant reduction in injuries. However, there was no statistically significant reduction in fatalities when a state has proportionally more interstate lane miles. Furthermore, the results indicated that states with a larger share of arterial lane miles in their networks have more fatalities and injuries, and those with more collectors have more injuries.

The National Roads Authority of Ireland has conducted a study with the main objectives being identifying High Accident Locations (HALs) on the inter-urban national road network, and investigating whether at each of these individual sites there were any risk factors, such as skidding etc, that may have influenced the number of accidents occurring at that section. The report of the study introduced Road Traffic Accidents (RTAs) as rare, random, multi-factor events in which a road user/s fails to cope with the surrounding traffic environment. The occurrence of RTA is by its very nature unpredictable. However, according to the report, while each individual accident is fundamentally unpredictable by its nature, the number of accidents in given locations over given time periods may display notable patterns and can be influenced by a range of factors. The report indicates that Multiple Linear Regression Models, Poisson Models, Negative Binomial Models and Accident Rate Models can be used in analyzing RTA. It is further discussed in the report that researchers have gradually moved away from the use of Multiple Linear Regression Models as a result of well-documented statistical difficulties. The Poisson distribution which is a discrete distribution with the variable in question taking on whole number values greater than or equal to 0, and which is often used to model the number of events occurring within a given period was considered for modeling purpose. The core aspect of the study was to develop a statistical model which, for inter-urban national roads, generates the expected number of accidents in a section given its traffic volume and section length. In circumstances where the number of recorded accidents significantly exceeds the expected number, then that section is considered as a HAL. The basic or null hypothesis studied was that sections of road of equal length with equal vehicle miles traveled should have similar levels of accidents over a given time period; and that the number of accidents in a given period should be similar to that estimated using the statistical techniques outlined above. The null hypothesis was rejected if the difference between the actual and estimated number was so large as to have arisen in less than one in twenty times. In other words, if it can be said to a ninety-five per cent confidence level that there is a difference between the actual and expected number of accidents then the null hypothesis is rejected in favor of the alternate hypothesis – there is a difference (NRA, 2007**).**

Tessma et al (2005), studied injury severity levels resulting from an accident using real data obtained from the Addis Ababa traffic office. Their research was focused on developing adaptive regression trees to build a decision support system to handle road traffic accident analysis for Addis Ababa city traffic office. As stated in their study, the Classification and Adaptive Regression Trees (CART) methodology is technically known as binary recursive partitioning. The process is binary because parent nodes are always split into exactly two child nodes and recursive because the process can be repeated by treating each child node as a parent. The key elements of a CART analysis are a set of rules for splitting each node in a tree: a) deciding when a tree is complete and b) assigning each terminal node to a class outcome (or predicted value for regression). They further stated that the specific attributes by which a given accident can be described are date and time, accident id , driver's name, vehicle type, driver's age, driver's gender, driver's educational level, driver's license status, relation of the driver and vehicle, driver's experience, possession of the vehicle, vehicle defect, vehicle age, accident area, accident road name, road segment separation, road direction, road surface type, roadway surface condition, light condition, weather condition, vehicle maneuver, accident type, total vehicles involved, total number of victims, accident victims category, victims profession, victims health condition, pedestrian maneuver, vehicle plate number, cost estimate of the damage and cause for accident. In addition to the input variables mentioned above the output variable for the research that is injury severity was also another attribute of a given accident. The target attribute, injury severity, has four classes: fatal injury, property damage (no injury), serious injury and slight injury. In Defining the Data Mining Function, they stated that each individual accident record in the data set is an input/output pair with each record having an associated output. The output variable, the injury severity, is categorical and as described above, has four classes. After successive experiments in building the best decision tree model, the next step of their study was to generate rules by tracing through the branches up to leafs. A rule is a correlation found between the main variable (dependent) and the others (independent). From the general rule, they found out that it was easy to see that, there is a 7.8 % chance that Injury-Severity will be a fatal injury, a 67.63% chance that Injury-Severity will be partial damage, a 8.15% chance that Injury-Severity will be serious injury and a 16.39% chance that Injury-Severity will be slight injury and this reveals that about 33% of the accidents, results in injury of different level half of which is either fatal or serious injury. Further, their study indicated that accident types involving pedestrians and single vehicle turn over mostly results in either fatal or serious injuries. Denying pedestrian priority and over speeding were also the top most determinant factors in injury severity. It is also apparent, as per their study, that if accident cause is driving with alcohol and accident type is vehicle-peds there is high probability of an accident resulting in fatalities or injuries. Over speeding is also another important factor especially if associated with vehicle-peds type of accidents. In general, the rules presented above indicate the possible conditions in which an accident will result in either of the injury severity classes. Moreover, the rules generated have indicated that attributes such as *'accident cause'*, *'accident type'*, *'driver age'*,*' road surface type'*, *'road condition'*, *'vehicle type'*, and *'light condition'* are found to be important variables for classification of accident severity. The study also indicated that, these variables are playing a significant role in all experiments being placed at the higher level of the tree which indicates their statistical significance than other variables like *'sex'*, *'weather condition'* and *'accident_id'*. Decision of selecting the best decision tree was based on the soundness of the rules generated as well as the number of misclassified records of different levels of injury severity.

Soyoung Jung et al (2009), assessed the effects of rainfall on the severity of single-vehicle crashes on Wisconsin interstate highways utilizing polychotomous response models. Weather-related factors considered in this study include estimated rainfall intensity for 15 min prior to a crash occurrence, water film depth, temperature, wind speed/direction, stopping sight distance and deficiency of car-following distance at the crash moment. For locations with unknown weather information, data were interpolated using the inverse squared distance method. Non-weather factors such as road geometrics, traffic conditions, collision types, vehicle types, and driver and temporal attributes were also considered. Two types of polychotomous response models were compared: ordinal logistic and sequential logistic regressions. The sequential logistic regression was tested with forward and backward formats. Comparative models were also developed for single vehicle crash severity during clear weather. In conclusion, the backward sequential logistic regression model produced the best results for predicting crash severities in rainy weather where rainfall intensity, wind speed, roadway terrain, driver's gender, and safety belt were found to be statistically significant. Their study also found that the seasonal factor was significant in clear weather. The seasonal factor is a predictor suggesting that inclement weather may affect crash severity. These findings can be used to determine the probabilities of single vehicle crash severity in rainy weather and provide quantitative support on improving road weather safety via weather warning systems, highway facility improvements, and speed limit management.

An assessment of 1,508 motorcycle accidents in Australia was done by Williams and Hoffmann in 1974. Data show that 63.7% of all accidents occurred during hours during day, 29.0% during night, and 7.3% at dawn and dusk. This is because the flow of traffic during day is greater than that of dawn and dusk contrary to the later findings of Yau (2003). Yau (2003) considered five factors, namely, the month of occurrence of the accident, day of the week, time of the accident, street lighting conditions and weather conditions. Results revealed that the day of the week and time of the accident are important factors affecting injury severity. It was also found that motorcycle accidents have higher risk during weekends than weekdays because motorcycle drivers driving on weekends may exhibit risky driving behaviors. Furthermore, motorcycle accidents are also more likely to happen between night and midnight. Perhaps during this time, the driver is more susceptible to driving at fast speeds due to the small number of vehicles on the road.

Savolainen and Mannering (2007) predicted motorcyclists' injury severities in single and multiple crashes using nested logit and multinomial logit models, respectively. Wet pavement was significant to increase no injury severity only in single-vehicle motorcycle crashes while none of the weather related factors were found to be significant to motorcyclists' injury severities in multi-vehicle crashes involving motorcyclists. Rainfall-related effect on crash severity outcomes has been also identified along with roadway characteristics.

Ordered probit models were used in a Abdel-Aty's study (2003) to predict driver injury severity in Central Florida, with crashes occurring in specific roadway sections, signalized intersections and toll plazas in expressway systems. It was found that crashes happening in signalized intersections with bad weather and dark street lighting had a significantly higher probability of severe injury. Abdel-Aty (2003) stated that an angle and turning collision in the adverse weather and dark street light conditions was a possible reason to contribute higher probability of injuries in signalized intersections.

Donnell and Mason (2004) employed ordinal and nominal logistic regression models to predict interstate highway crash severity in cross-median and median-barrier collisions, respectively. Researchers found wet or icy pavement surface to be significant factor in decreasing crash severity.

In contrast, in a study by Lee and Mannering (2002), the nested logit analysis showed that wet roadway surfaces increased the likelihood of evident and disabling injury/fatality in run-off-roadway accidents. The conflicting results from these two studies suggest the need for a detailed analysis for weather-related crash severity by collision type.

Jovanis and Chang (1986) found a number of problems with the use of linear regression in their study applying Poisson regression as a means to predict accidents. For example, they discovered that as vehicle-kilometers traveled increases, so does the variance of the accident frequency. Thus, this analysis violates the homoscedasticity assumption of linear regression. In a well-summarized review of models predicting accident frequency, Milton and Mannering (1997) state that "the use of linear regression models is inappropriate for making probabilistic statements about the occurrences of vehicle accidents on the road". They showed that the negative binomial regression is a powerful predictive tool and one that should be increasingly applied in future accident frequency studies. Nassar et al. (1997) developed an integrated Accident Risk Model (ARM) for policy decisions using risk factors affecting both accident occurrences on road sections, and injury severity of occupants involved in the accidents. Using negative binomial regression and a sequential binary logit formulation, the models they developed are practical and easy to use. Mercier et al. (1997) used logistic regression to determine whether either age or gender (or both) was a factor influencing severity of injuries suffered in head-on automobile collisions on rural highways. Logistic regression was also used by Veilahti et al. (1989) in predicting automobile driving accidents of young drivers. They examined the predictive values of the Cattel 16-factor personality on the occurrence of automobile accidents among conscripts during 11-month military service in a transportation section of the Finnish Defense Forces.

James and Kim (1991) developed a logistics regression model for describing the use of child safety seats for children involved in crashes in Hawaii from 1986 through 1991. The model reveals that children riding in automobiles are less likely to be restrained; drivers that use seat belts are far more likely to restrain their children; and one-and-two-year olds are less likely to be restrained.

On the issue of age, the study of Abdel-Aty and Radwan (2000) reveals that young and older drivers have a larger possibility of accident involvement than middle-aged drivers when experiencing heavy traffic volume do. Young drivers are also subject to involvement in speeding, increasing the risk of getting into an accident. Data used in the study was taken from Central Florida, U.S.A. This result may not be the same for other studies due to the difference in context and proportion of people in certain driver age groups. As of 2000, sixty-five percent of the population was part of the middle-age group in the Philippines. Hence, age is a factor that should be considered in the occurrence of crashes. However, it is evident in the negative binomial model used in Abdel-Aty and Radwan (2000) that the factors were correlated. With this, the results may not reflect the actual relationships of each factor in the study done.

Previous studies presented only considered single factors on accident occurrence or injury severity. On the other hand, it can be possible that different factors as well as their interactions are also considered. The study of Yannis et al. (2005) and Al-Ghamdi (2002) conducted in Greece and Saudi Arabia, respectively, considered the combined effects of factors including age on different response variables. The former tackles on combined effects of driver and motorcycle dependent factors on two response variables namely accident severity and at-fault risk. It has been found that there is a significant combined effect of age and engine size with respect to accidents. On another context, Al- Ghamdi (2002) had used logistic regression in developing a model that determines the relationships between nine variables namely, location, accident type, collision type, accident time, accident cause, driver age, nationality, vehicle type, and license status. The relationships of the variables in terms of correlations of the factors are analyzed through the use of an odds matrix. However with this approach, only two variables are found to be significant and thus included in the model. It has been found that there is a relationship between accident location and accident cause, which can help in the development of specific awareness programs. For example, the odds of being involved in an accident at a non-intersection location due to going the wrong way violation is higher than for any other violation, drivers can be warned about the lethality of the violation. Results for both studies show that interactions between factors significantly affect the response variables.

Compared with previous studies, this study applied logistic regression to accident records from MTTU's Northern Regional Office taken at crash moment to predict crash severity outcomes. Research findings from this study will provide guidance on counter measures to prevent severe crashes and improve overall safety.

CHAPTER 3

METHODOLOGY

3.0 INTRODUCTION

This Chapter deals with the method of data collection and how the data is organized for the application logistic regression analysis. It also contains detailed discussion of the main statistical techniques and a review of the basic methods used for the analysis of the data.

3.1 DATA SOURCE

Accidents are recorded by the motor traffic and transport unit of the police on daily basis. The data for this project were solely secondary data which was taken from Motor Traffic and Transport Unit of the Ghana Police Service, Northern regional office, Tamale. Three years of data, containing information on accident - involved drivers for the period of 2007 to 2009, was used in this study. The data contained those fatal accidents that had occurred in the region within the period under consideration.

3.2 REVIEW OF BASIC METHODS FOR ANALYSIS - LOGISTIC REGRESSION ANALYSIS

Logistic regression is useful for predicting a binary dependent variable as a function of predictor variables. The goal of logistic regression is to identify the best fitting model that describes the relationship between a binary dependent variable and a set of independent or explanatory variables. The dependent variable is the population proportion or probability, P, that the resulting outcome is equal to one. Parameters obtained for the independent variables can be used to estimate odds ratios for each of the independent variables in the model.

3.2.1 THEORETICAL BACKGROUND OF LOGISTIC REGRESSION

It is important to understand that the goal of an analysis using logistic regression is the same as that of any model-building technique used in statistics: To find the best fit and most parsimonious. What distinguishes a logistic regression model from a linear regression model is the outcome variable. In the logistic regression model, the outcome variable is binary or dichotomous. The difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. Once this difference is accounted for, the methods employed in an analysis using logistic regression follow the same general principles used in linear regression analysis. In any regression analysis the key quantity is the mean values of the outcome variable, given the values of the independent variable as:

$$\mathrm{E}(Y/x) = \beta_0 + \beta_1 x \qquad (1)$$

Where Y denotes the outcome variable, x denotes a value of the independent variable, and the β_i's denote the model parameters. The quantity is called the conditional mean or the expected value of Y given the value x. Many distribution functions have been proposed for use in the analysis of a dichotomous outcome variable. The specific form of the logistic regression model is

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \qquad (2)$$

Where, for simplifying notations, we let $\pi(x) = E(Y/x)$. The transformation of $\pi(x)$ logistic function is known as the logit transformation:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x \qquad (3)$$

The importance of this transformation is that $g(x)$ has many of the desirable properties of a linear regression model. The logit, $g(x)$ is linear in its parameters, may be continuous, and may range from minus infinity to positive infinity, dependent on the range of x.

In summary the main features in a regression analysis when the outcome variable is dichotomous as follows:

1. The conditional mean of the regression equation must be formulated to be bounded between 0 and 1 (equation (2) satisfies this constraint).

2. The binomial, not the normal, distribution describes the distribution of the errors and will be the statistical distribution upon which the analysis is based.

3. The principles that guide an analysis using linear regression will also derive for logistic regression.

In linear regression the method used most often for estimating unknown parameters is least squares. In that method the values of parameters are chosen to minimize the sum of squared deviations of the observed values of Y from the modeled values. Under the assumptions for linear regression the method of least squares yields estimators with a number of desirable statistical properties. Unfortunately, when the method of least squares is applied to a model with a dichotomous outcome the estimators no longer have these same properties. The general method of estimation that leads to the least squares function under the linear regression model (when the error is normally distributed) is called the maximum likelihood. This method provides the foundation for estimating the parameters of a logistic regression model. A brief review of fitting the logistic regression model is given below. For further details one can read Hosmer and Lemeshow (1989).

If Y is coded as zero or one (binary variable) then the expression $\pi(x)$ given in equation (2) provides the conditional probability that Y is equal to 1 given x. This will be denoted as $P(Y=1\backslash x)$. It follows that the quantity $1-\pi(x)$ gives the conditional probability that Y is equal to zero given x, $P(Y=0\backslash x)$. Thus, for those pairs (x_i, y_i), where $y_i = 1$ the contribution to the likelihood function is $\pi(x_i)$, and for those pairs where $y_i = 0$ the contribution to the likelihood function is $1-\pi(x_i)$, where the quantity denotes the $\pi(x_i)$ values of $\pi(x)$ computed at x_i. A convenient way to express the contribution to the likelihood function for the pair (x_i, y_i) is through the term:

$$\zeta(x_i) = \pi(x_i)^{y_i}\left[1-\pi(x_i)\right]^{1-y_i}$$

Since x_i's are assumed to be independent, the product for the terms given in the above equation gives the likelihood function as follows:

$$l(\beta) = \prod_{i=1}^{n} \zeta(x_i) \qquad (4)$$

It is easier mathematically to work with log of equation (3) which gives the log likelihood expression:

$$L(\beta) = \ln(\left[l(\beta)\right] = \sum_{i=1}^{n} \left\{y_i \ln\left[\pi(x_i)\right] + (1-y_i)\ln\left[1-\pi(x_i)\right]\right\} \qquad (4.1)$$

Maximizing the above function with respect β to and setting the resulting expressions equal to zero will produce the values of β as follows:

$$\sum_{i-1}^{n}\left[y_i - \pi(x_i)\right] = 0 \qquad (5)$$

and

$$\sum_{i-1}^{n} x_i \left[y_i - \pi(x_i) \right] = 0 \qquad (6)$$

These expressions are called likelihood equations. An interesting consequence of equation (5) is that:

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{\pi}(x_i)$$

That is, the sum of the observed values of y is equal to the sum of the expected (predicted) values. This property is especially useful in assessing the fit of the model (Hosmer and Lemeshow, 1989).

After estimating the coefficients, the assessment of the significance of the variables in the model is taking place. If y_i denotes the observed value and $\hat{y}_i$ denotes the predicted value for the *ith* individual under the model, then the statistics used in the linear regression is:

$$\text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

The change in the values of SSE is due to the regression source of variability, denoted

$$\text{SSR} = \left[\sum_{i=1}^{n} (y_i - \overline{y_i})^2 \right] - \left[\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \right]$$

Where $\overline{y}$ is the mean of the response variable. Thus, in linear regression, interest focuses on the size of R^2. A large value suggests that the independent variable is important, whereas a small value suggests that the independent variable is not useful in explaining the variability in the response variable.

The principle in logistic regression is the same. That is, observed values of the response variable should be compared to the predicted values obtained from models with and without the variable in question. In logistic regression this comparison is based on the log likelihood function defined in equation (4.1). Defining the saturation model as one that contains as many parameters as there are data points, the current model is the one that contains only the variable under question. The following ratio:

$$D = -2\ln\left[\frac{likelihood \quad of \quad the \quad current \quad \mathrm{mod}\,el}{likelihood \quad of \quad the \quad saturated \quad \mathrm{mod}\,el}\right] \qquad (7)$$

is called the likelihood ratio. The reason for the minus two is that its log is mathematical. Using equation (4.1) and equation (7), the following test statistic can be obtained:

$$D = -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\hat{\pi}_i}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\hat{\pi}_i}{1-y_i}\right)\right] \qquad (8)$$

Where $\hat{\pi}_i = \hat{\pi}(x_i)$

The statistics D, in equation (8) is called, at least in this study, deviance and plays an essential role in some approaches to the assessment of goodness-of-fit. The deviance for logistic regression plays the same role that the residual sum of squares plays in linear regression (i.e., identically equal to SSE).

For the purpose of assessing the significance of an independent variable, the value of D should be compared with and without the independent variable in the model. The change in D, due to inclusion of the independent variable in the model, obtained as follows:

G = D(for the model without the variable) – D(for the model with the variable)

This statistic plays the same role in logistic regression as does the numerator of the partial F test in linear regression. Because the likelihood of the saturated model is common to both values of D being the difference to compute G, it can be expressed as:

$$G = \left[\frac{likelihood \quad without \quad the \quad \mathrm{var}\,iable}{likelihood \quad with \quad the \quad \mathrm{var}\,iable}\right] \qquad (9)$$

It is not appropriate here to derive the mathematical expression of the statistic G. Yet, it should be said that under the null hypothesis that is β_1 equal to zero, G will follow a chi-square distribution with 1 d.f. Another test statistic, similar to G for the purpose, used in this study is known as Wald Statistic (W) which follows a standard normal distribution under the null hypothesis that $\beta_1 = 0$. This statistic is computed by dividing the estimated value of the parameter by its standard error as:

$$W = \frac{\hat{\beta}_1}{S\hat{E}\left(\hat{\beta}_1\right)} \tag{10}$$

It should be mentioned that Wald test sometimes fail to reject when the coefficient was significant, and hence, the likelihood ratio test should be used in suspicious cases.

The odds ratio (OR) is a ratio of two odds. An odds ratio can be used to give us an idea of how strongly a given variable may be associated with the outcome of interest compared to other variables. For a probability of success *p*, the odds (likelihood) of success (in our case with risk factors, i.e., fatality) are defined as shown in Equation 11:

$$odds = \frac{\pi}{1-\pi} \tag{11}$$

Essentially, odds are nonnegative values. When the odds are less than one, the probability of success is less than that of failure; when the odds equal one, the probabilities of success and failure are equally likely; and when the odds are greater than one, the probability of success is greater than that of failure. In a nutshell, the odds ratio compares the odds of success proportion (fatality) for Group 1 (e.g., female) to the odds of success proportion (fatality) for Group 2 (e.g., male) when analyzing a variable of interest, e.g., GENDER in this case. Generally, the odds ratio compares the odds of success proportion in row 1 (Group 1) to the odds of success proportion in row 2 (Group 2) of a 2 x 2 contingency table (see Table 3.2) and calculated as shown in Equation 12.

$$\text{OR} = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} \qquad (12)$$

Table 3.2 Contingency table describing the formulation of odds ratio formula

INJURY TYPE (binary options for dependent variable	Independent variable: GENDER	
	Probabilities of the binary options	
	Group 1 (Female)	Group 2 (Male)
Fatality (Success)	π_1	π_2
Non-Fatality	$1-\pi_1$	$1-\pi_2$

An odds ratio of one indicates that the success or desired event under study is equally likely in both groups. An odds ratio greater than one indicates that the successful event is more likely to occur in the first group than in the second group and an odds ratio less than one indicate that the successful event is less likely in the first group. The farther odds ratio values are from 1.0 in either direction, the stronger the association among the variables (Agresti, 2007).

3.3 DESCRIPTION OF VARIABLES IN THE ANALYSIS

The following variables are considered in this study.

3.3.1 DEPENDENT VARIABLE

Severity of accident is a variable in the MTTU's datasets recorded as 'consequences of accident' and is assigned to all people involved in a traffic crash and it describes the injury severity level each person sustained when a traffic crash occurred. In the MTTU's datasets, the variable 'consequences of accident' is coded with the following options:

1. Number of minor injuries

2. Number of serious injuries

3. Number of fatalities

This variable was recorded as a binary variable as follows:

0 if the accident results in at least one injury but no fatality, and 1 when there is at least one fatality resulting from the accident.

3.3.2 INDEPENDENT VARIABLES

1. Gender – this variable reports the gender of the person who got involved in the accident and since this is a binary variable it was left unchanged with its original responses.

- Female "1"
- Male "0"

2. Age – this variable represents the age of the person involved in the accident.

3. Date of accident – the date of the accident was recorded.

4. Use of alcohol when accident occurred was recorded as;

- Yes “1”
- No “0”

5. Use of safety belt when accident occurred was recorded as;

- Yes “1”
- No “0”

6. Vehicle ownership - the categories are:

- Owner
- Hired
- Other (which may be relative, friend, one who rented the vehicle …etc)

7. Type of vehicle;

- Private (salon cars, small cars, ets) “1”
- Commercial “0”

8. Age of vehicle – This is the number of years since the vehicle was manufactured.

9. Weight of vehicle;

- Light weight “1”
- Heavy duty “0”

10. Vehicles tyres condition – the categories are:

- New or good as new
- Somewhat worn
- Quite worn
- Very worn

11. Estimated Speed of vehicle when accident occurred;

- 50km/h and below “1”
- Above 50km/h “0”

12. Posted speed limit at the site of accident;

- 50km/h “1”
- Others “0”

13. Road surface condition when accident occurred;

- Wet Surface “1”
- Dry Surface “0”

14. Weather condition when accident occurred – this variable describes the weather conditions at the time of the accident. It was recorded as a binary variable as follows:

- Good Weather “1”.

- Other: it includes cloudy, rain, snow, and other “0”.

15. Traffic lighting condition when accident occurred;

- Red
- Yellow
- Green

“1” was assign to the option which was applicable and “0” for the rest.

16. Drivers familiarity of route.

- Pass scene of accident at least once a month “1”.
- More seldom pass than once in a month “0”

17. Type of driving license;

- License A
- License B
- License C
- License D
- License E
- License F

“1” was assign to the option which was applicable and “0” for the rest.

18. Age of driving license - This is the number of years since the driver received a driving license.

19. Scene of accident was recorded as a nominal variable with the following categories are:

- Junction
- link
- other

20. Cause of accident – the categories from the MTTU's questionnaires are listed below. For the purpose of the research which concentrates on identifying the factors that relates to severity of road accidents, each of the under listed cause of accident was recorded as a variable. 1 was assigned to a variable which contributed in causing the accident and 0 otherwise.

- Excessive speeding
- Inattention, confusion of lack of judgment of driver
- Drivers careless at road junction and cutting corners
- Improperly overtaking or cutting in
- Inexperience of driver
- Intoxication
- Other recklessness or negligence by drivers
- Over loading

- Mechanical defects
- Defective lights
- Dazzling lights
- Skid and road surface defects
- Other road defects
- Obstructions
- Level erosions
- Children
- Adults crossing road carelessly
- Adults boarding or alighting from vehicles
- Other pedestrian faults
- Passengers faults
- Animals not under control
- Recklessness or negligence caused by peddle cyclist
- Recklessness or negligence caused by drivers of horse - driven vehicles

3.4 MODEL DESCRIPTION

The dependent variable in this research is Accident and of dichotomous type and stands for accident severity. Each accident in the sampled data was categorized as either non-fatal or fatal. The logistic model used is:

$$P(Non-Fatal) = \pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}}$$

And thus

$$P(Fatal) = 1 - P(Non-Fatal) = 1 - \pi(x) = 1 - \frac{e^{g(x)}}{1+e^{g(x)}} = \frac{1}{1+e^{g(x)}}$$

Where *g(x)* stands for the function of the independent variables as:

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$$

Logistic regression determines the coefficients that make the observed outcome (Non-fatal or Fatal Accident) most likely using the maximum-likelihood technique. The independent variables could be continuous or dichotomous. For the latter, there should be special coding with the use of dummy variables. These dummy variables should be defined in a manner consistent to the SPSS software used in this study (SPSS 19). The Wald tests, together with the Deviance, will be used as criteria to include or remove independent variables in the model. The SPSS software has built in routines to obtain Deviance and estimates of the model parameters.

CHAPTER 4

DATA ANALYSIS AND RESULTS

4.0 INTRODUCTION

TABLE 4.1 below is a snapshot of the data used for this project which is made up of 398 accident records from MTTU northern region. Detailed results of the analysis can be found in appendix B. The results of this study give an outlook of risk factors which might be possible causes of fatalities in accidents in the Northern Region and can be isolated when attempting to prevent fatal injuries when accidents occur.

Table 4.0 Accident records from MTTU Northern Region (2007-2009)

Fatality	Gender	Age	Date	Alcohol	Safetybelt	Vownership
1	0	34	03-Dec-2009	0	0	1
1	0	55	26-Dec-2009	1	0	1
1	0	40	25-Dec-2009	0	0	2
1	0	52	11-Dec-2009	1	0	1
1	0	27	05-Dec-2009	0	0	1
1	0	25	13-Dec-2009	0	0	2
1	0	26	17-Dec-2009	0	0	2
1	0	28	26-Dec-2009	1	0	1
1	0	26	14-Dec-2009	1	0	1
1	0	24	21-Dec-2009	1	0	3
1	0	33	04-Dec-2009	1	0	1
1	0	25	29-Dec-	1	1	1

			2009			
0	0	30	24-Dec-2009	0	0	1
0	0	55	23-Dec-2009	1	0	1

4.1 Descriptive Statistics

A review of the accidents data revealed that among the 398 records, 3.1% involved minor injuries. The rest were considered as accidents with fatal injuries. The wide gap between the percentages is due to the fact that most accidents are not officially reported to the police especially those without fatalities.

Table 4.1 Consequences of accident

	Frequency	Percent	Valid Percent	Cumulative Percent
No fatality	5	1.3	1.3	1.3
Fatality	393	98.7	98.7	100.0
Total	398	100.0	100.0	

4.1.1 Age of Drivers involved in the accidents

The driver's age was grouped into three categories. Among these categories, drivers with in the age group 26-50 are responsible for the larger number of accidents (55.8%). Drivers with age 50+ have the smallest in all the three measurements (8.8%) and 35.4% for drivers with age group 25 maximum.

Table 4.2 Age of drivers involved in Accidents

	Frequency	Percent	Valid Percent	Cumulative Percent
25 maximum	141	35.4	35.4	35.4
26-50	222	55.8	55.8	91.2

50+	35	8.8	8.8	100.0
Total	398	100.0	100.0	

4.1.2 Gender of Drivers involved in accidents

Out of the 398 records, males constituted 91.2% of all the accidents in the northern region in the period 2007-2009 whiles females constituted 8.8%.

Table 4.3 Gender of person involved in accident

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	363	91.2	91.2	91.2
Female	35	8.8	8.8	100.0
Total	398	100.0	100.0	

Furthermore, 113 males and 28 females were in the age group of 25 maximum, 216 males and 6 females in the 26-50 age group and 34 males and 1 female in the 50+ age group.

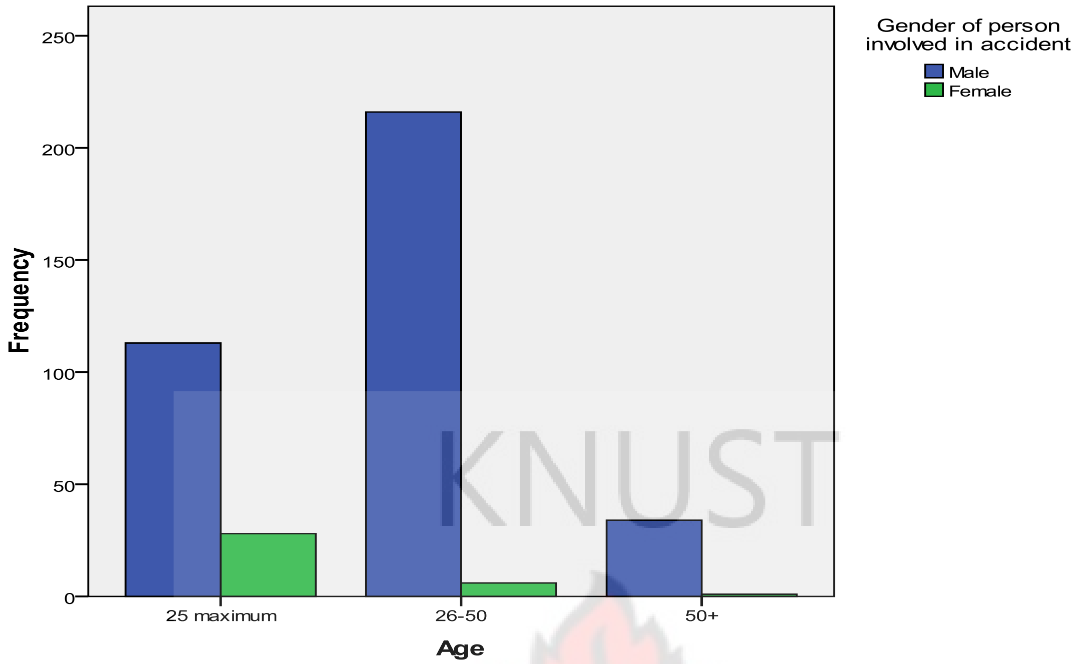

Figure 4.1: Bar chart showing Gender of persons involved in accidents and Age

4.1.3 Use of Alcohol

Out of the 398 records analyzed, 214 (53.8%) of them drove under the influence of alcohol while 184 (46.2%) drove normal, that is not under the influence of alcohol.

Table 4.4 Use of Alcohol

	Frequency	Percent	Valid Percent	Cumulative Percent
No	184	46.2	46.2	46.2
Yes	214	53.8	53.8	100.0
Total	398	100.0	100.0	

Also, out of the 214 people who drove under the influence of alcohol, 211 (98.6%) were males whiles 3 (1.4%) were females.

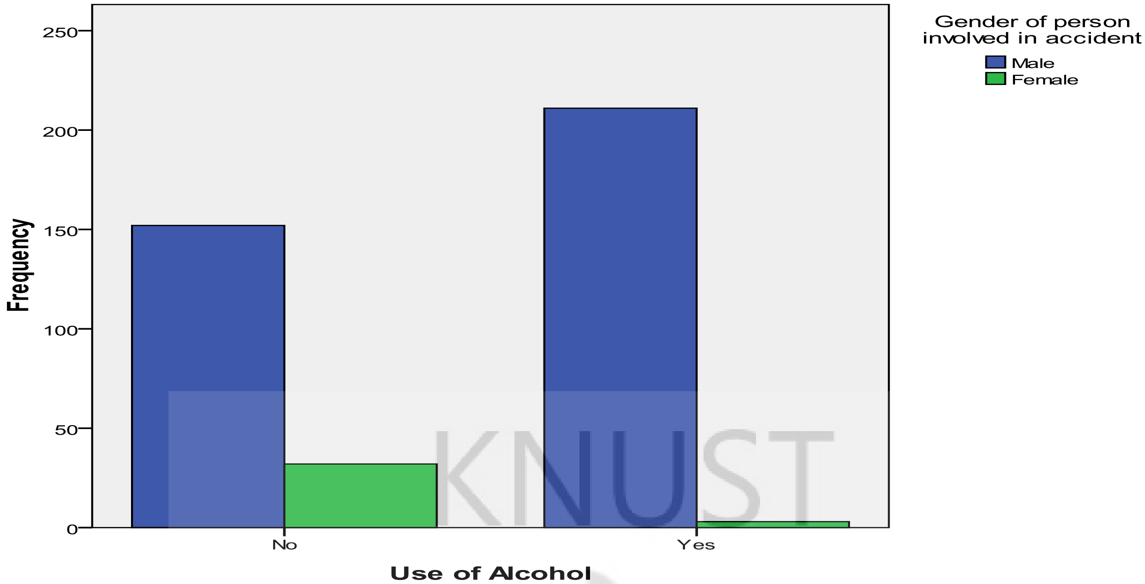

Figure 4.2: Bar chart showing use of alcohol and gender

4.1.4 Use of Safety Belt

The results revealed that 316 people representing 79.4% did not use safety belts whiles driving and 82 people representing 20.6% drove whiles using safety belts.

Table 4.5 Use of safety belt

	Frequency	Percent	Valid Percent	Cumulative Percent
No	316	79.4	79.4	79.4
Yes	82	20.6	20.6	100.0
Total	398	100.0	100.0	

4.1.5 Driving Experience

When Driving Experience is considered (number of years of driving license), drivers with experience less than 5 years were 53 (13.3%), 5-10 years of experience were 213 (53.5%) and above 10 years of experience were 132 (33.2%). It is worth mentioning that drivers who had no diving license were 9 out of the total 398 representing 2.3%.

Table 4.6 Driving Experience of persons involved in accident

	Frequency	Percent	Valid Percent	Cumulative Percent
less than 5 yrs	53	13.3	13.3	13.3
5-10	213	53.5	53.5	66.8
11 and above	132	33.2	33.2	100.0
Total	398	100.0	100.0	

Also, the figure below shows an analysis of the age group and driving experience of persons involved in the accidents.

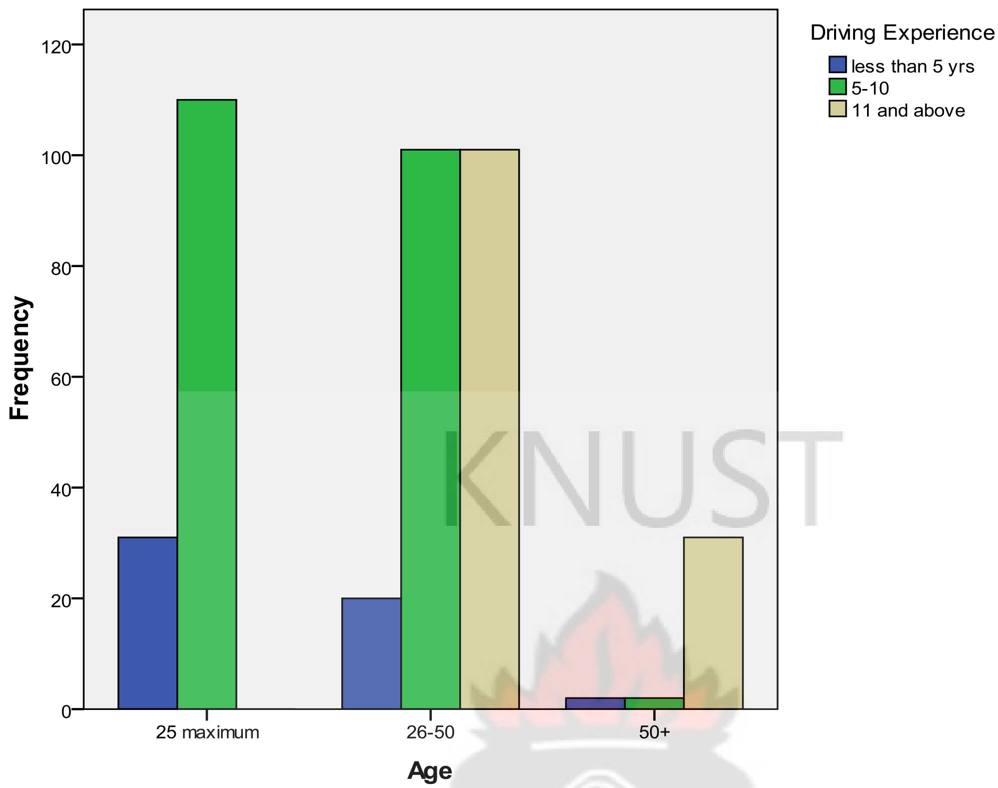

Figure 4.3 Bar chart showing Driving Experience and Age of Drivers

4.2 Variable Selection

The analysis was started by testing the significance of the association each explanatory variable could have with the dependent variable. For this purpose the entering selection process of logistic regression was followed in this study. First, all of the variables with no interactions were tested based on Wald (W) statistic. The goal was to eliminate, at the beginning, those variables that were not significant and then continue with testing interaction effects with only significant variables. Table 4.7 presents the results from fitting all the explanatory variables simultaneously.

Table 4.7 Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	18.736	5787.218	.000	1	.997	1.371E8
Age	-.014	.052	.073	1	.787	.986
EsiveSpeed	-2.395	2.110	1.289	1	.256	.091
Inattentn	13.839	9640.361	.000	1	.999	1023837.090
CarelessJ	14.117	7159.984	.000	1	.998	1351482.719
WOvertaking	-2.911	1.929	2.278	1	.131	.054
Inexperience	-3.633	1.903	3.643	1	.056	.026
Intoxication	14.474	4116.495	.000	1	.997	1931722.343
OReckIness	14.226	4631.994	.000	1	.998	1507534.975
OverLoading	-6.407	2.413	7.049	1	.008	.002
MechDefects	-3.608	2.021	3.186	1	.074	.027
LightsDefects	17.592	40192.969	.000	1	1.000	4.365E7
RSurfDefects	-4.696	28997.819	.000	1	1.000	.009
ORDefects	13.694	19037.554	.000	1	.999	885773.331
Obstructn	-4.851	2.390	4.121	1	.042	.008
Children	14.160	28373.733	.000	1	1.000	1410780.912
AdultsCRC	13.140	11290.749	.000	1	.999	508803.884
OPedestrian	14.054	40192.970	.000	1	1.000	1268914.865
PassengerF	-18.806	57135.291	.000	1	1.000	.000
PeddleCyclist	13.467	21569.898	.000	1	1.000	706022.050
Constant	7.567	2.350	10.370	1	.001	1934.147

From the table above, it appears that the variables OverLoading and Obstructn show some significant effect whiles the rest of the variables are not significant. Therefore, further analyses were made only on those variables which showed some significant association with the dependent variable, but only two of the variables were found statistically significant at 0.05 levels.

Table 4.8 Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
WOvertaking	-2.587	1.430	3.273	1	.070	.075
OverLoading	-5.072	1.582	10.277	1	.001	.006
MechDefects	-2.154	1.425	2.286	1	.131	.116
Obstructn	-3.367	1.447	5.415	1	.020	.034
Constant	5.765	1.002	33.134	1	.000	319.000

Finally, analyses were made on the two variables that showed significant effects on the dependent variable and they still remain significant as shown in TABLE 4.9 below.

Table 4.9 Variables in the Equation

							95% C.I.for EXP(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
OverLoading	-4.148	1.355	9.373	1	.002	.016	.001	.225
Obstructn	-2.444	1.195	4.185	1	.041	.087	.008	.903
Constant	4.842	.580	69.771	1	.000	126.667		

4.3 Logistic Regression Analysis

As mentioned in section 3.4, Logistic Regression is the selected method for analyzing the accident data.

Only the explanatory variables found significantly associated with the dependant variable as per the Wald's test result are considered in the Logistic regression analysis. Accordingly two variables, namely OverLoading and Obstruction are included in the model.

4.3.1 The Logit Model

According to the analysis above, the logit model with the significant variables is:

$$\hat{g}(x) = 4.842 - 2.444 Obstructn - 4.148 OverLoading$$

Hence the logistic regression model developed in this study is:

$$P(Non - Fatal) \; = \; \pi(x) = \frac{e^{4.842 - 2.444 Obstructn - 4.148 OverLoading}}{1 + e^{4.842 - 2.444 Obstructn - 4.148 OverLoading}}$$

And thus

$$P(Fatal) \; = 1 - P(Non - Fatal) \; = \; 1 - \pi(x) = 1 - \frac{e^{4.842 - 2.444 Obstructn - 4.148 OverLoading}}{1 + e^{4..842 - 2.444 Obstructn - 4.148 OverLoading}}$$

4.3.2 Testing for the Significance of the Model

Once the model has been fit, the process of assessment of the model begins. Several tests, including Pearson chi-square and deviance, Wald Statistic, and the Hosmer-Lemeshow Tests, can be used to determine how effective the model is in describing the outcome variable. This is referred as its goodness-of-fit. These tests end up with a chi-square criterion to make the decision on the model fit. A very good reference for the theory of such tests is, for example, Hosmer and Lemeshow (1989). The validity of the model in this study was first checked by examining the statistical level of significance for its coefficients using Wald statistic as shown in TABLE 4.9. The model chi-square has a value of 7.931 and a probability of $p < 0.019$ indicating that the predictors are significant. Also table the classification shows that we are 98.7% right in the classification of our cases.

4.4 MODEL INTERPRETATION

The interpretation of any fitted model requires the ability to draw practical inferences from the estimated coefficients. The estimated coefficients, for the independent variables, represent the slope or rate of change of the dependent variable per unit of change in the independent variable. The slope coefficient in this model represents the change in the logit for a change of one unit in the independent variable *x*. Proper interpretation of the coefficient, in a logistic regression model, depends on being able to place meaning on the difference between two logits. The exponent of this difference (the difference between the two logits) gives the Odds Ratio which is defined as the ratio of the odds for the independent variable being present to the odds of not being present. The following subsections give illustrations for interpretation for the model developed in this study.

4.4.1 Impact of OverLoading on Accident Severity

The estimate for OverLoading is -4.184 can be transferred to the odds ratio as:

It should be said that OverLoading has two levels, 1 for being present and 0 for being absent. To interpret the parameter estimate (-4.184) for OverLoading, the logit difference should be computed as follows:

Logit (Fatal Accident/OverLoading (1)) = $4.842 - 2.444Obstructn - 4.14(1)$

Logit (Fatal Accident/OverLoading (0)) = $4.842 - 2.444Obstructn - 4.148(0)$

Logit difference = $4.842 - 2.444Obstructn - 4.148(1) - (4.842 - 2.444Obstructn - 4.148(0))$

=-4.148

Hence the odds ratio $e^{-4.148}$ = 0.016.

This indicates that the odds of being in a fatal accident with Overloading as a factor is 0.016 or 1.6%.

4.4.2 Impact of Obstructn on Accident Severity

The estimate for Obstructn is -2.444 can be transferred to the odds ratio as:

It should be said also that Obstructn has two levels, 1 for being present and 0 for being absent. To interpret the parameter estimate (-2.444) for Obstructn, the logit difference should be computed as follows:

Logit (Fatal Accident/Obstructn (1)) = $4.842 - 2.444(1) - 4.14 OverLoading$

Logit (Fatal Accident/Obstructn (0)) = $4.842 - 2.444(0) - 4.148 OverLoading$

Logit difference = $4.842 - 2.444(1) - 4.148 OverLoading$ - ($4.842 - 2.444(0) - 4.148 OverLoading$)

= -2.444

Hence the odds ratio $e^{-2.444}$ = 0.087

This indicates that the odds of being in a fatal accident with Obstructn as a factor is 0.087 or 8.7%. Details of the above can be seen in TABLE 4.9

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.0 INTRODUCTION

This thesis examined the factors relating to severity of accidents in the Northern Region during the period of 2007-2009. Since the response variable is of binary nature (i.e., has two categories: fatal or non-fatal), logistic regression technique was used to develop the model in this study. Excessive speeding, Inattention, confusion of lack of judgment of driver, Drivers careless at road junction and cutting corners, Improperly overtaking or cutting in, Inexperience of driver, Intoxication, Other recklessness or negligence by drivers, Over loading, Mechanical defects, Defective lights, Dazzling lights, Skid and road surface defects, Other road defects, Obstructions, Level erosions, Children, Adults crossing road carelessly, Adults boarding or alighting from vehicles, Other pedestrian faults, Passengers faults, Animals not under control, Recklessness or negligence caused by peddle cyclist, Recklessness or negligence caused by drivers of horse - driven vehicles were considered in this study to identify which of them contributed to accident fatalities in the Northern Region.

5.1 CONCLUSION

Using the concept of Deviance together with Wald Statistic, the study variables were subjected to statistical testing. Only two variables were included in the model, namely, OverLoading and Obstructn. The observed level of significance for regression coefficients for the two variables were less than 5% suggesting that these two variables were indeed good explanatory variables. The results also showed that the model provided a reasonable statistical fit.

Analysis of odds in this study also showed that accidents due to Obstruction are more likely to be fatal.

5.2 RECOMMENDATIONS

It can be learnt from this study that in addition to the efforts being made to reduce the frequency of accidents in general, special attention should be given to reduce the severity of accidents by taking the following into consideration:

- All accidents whether fatal or not should be reported to the police so as to update the police (MTTU) accident records.
- Strict control and management of vehicle movement is necessary especially in residential areas. The same would be required for areas around schools and religious institutions.
- Also, car accidents could be avoided by avoiding distraction in the car such as loud music, argument or the use of cell phones. Drivers should be aware that their full attention should be on roads so that they can make momentary decisions if ever they are confronted by a road condition.
- Drivers who are drunk should not drive. There should be more strict laws that would punish drunk drivers who risk their lives and that of other motorists and strict adherence to the load a vehicle can carry so as not to over load the vehicle and make it difficult to control.
- During rainy or unclear weather, drivers can reduce accidents by slowing down, leaving extra space between their car and the vehicle in front of them and keep their windshield clear.

- Additionally, vehicles should be equipped with air bags. Studies show that air bags alone can reduce the chance of fatal crash by 20 to 40 per cent and when air bag protection is combined with seat belt, the risk of a fatal crash can be reduced from 45 to 55 per cent.

- Further studies can be made on the area of road accidents by considering detail and accurate information on various variables. For example if the causes and consequences of an accident are recorded in detail instead of broad categories results could be more accurate and efficient.

Accidents happen everywhere, every day even to the most careful drivers. It is however, everyone's responsibility to prevent or avoid road accidents. Everyone in a "small way" can contribute effectively to reduce car accidents on our roads.

REFERENCES

1. Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered Probit models. *Journal of Safety Research*, 34 (5), 597–603.
2. Abdel-Aty, M.A. & Radwan, A.E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*, 32, 633-642.
3. Agresti, A. (2007). *An Introduction to categorical data analysis.* 2nd Edition, John Wiley and Sons, Inc., New York.
4. Akongbota J. (2011, May 12). Reducing accidents on our roads. Retrieved December 21, 2011, from government of Ghana official portal web site: http://www.ghana.gov.gh/index.php?option=com_content&view=article&id=5864:reducing-accidents-on-our-roads&catid=24:features&Itemid=167.
5. Al-Ghamdi, A.S. (2002). Using logistic regression to estimate the influence of accident factors on accident severity. *Accident Analysis and Prevention*, 34, 729-741.
6. Donnell, E.T., & Mason Jr., J.M. (2004). Predicting the severity of median-related crashes in Pennsylvania by using logistic regression. *Transportation Research Record*, 1897, 55–63.
7. Flores, G. M. T., Gotohio, M. P. T., Paras, N. G. C., & Seva, R. R. (2009). Analysis Motorcycle Accidents Based on Environmental and Personal Factors. Department of Industrial Engineering De La Salle University, Manila, Philippines.
8. Ghana news agency. (2010, May 21). Road accidents cost Ghana millions of dollars. Retrieved November 11, 2011, from Ghanaweb web site: http://ghanaweb.com/GhanaHomePage/NewsArchive/artikel.php?ID=182523.

9. Hilakivi, I., Veilahti, J., Asplund, P., Sinivuo, J., Laitinen, L., & Koskenvuo K. (1989). A sixteen-factor personality test for predicting automobile driving accidents of young drivers. *Accident Analysis and Prevention*, 21(5), 413-418.

10. Hosmer, D. W. & Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley & Sons, New York.

11. James, J. L., & Kim, K. E. (1986-1991). Restraint use by children involved in crashes in Hawaii. *Transportation Research Record*, 1560, 8-11, 19.

12. Jovanis, P., & Chang, H. (1986). Modeling the relationship of accidents to miles traveled. *Transportation Research Record,* 42-51.

13. Kwame Y. (2011, August 21). Congratulations GPRTU, but muslim drivers also drink. Retrieved December 21, 2011, from Ghanaweb web site: http://www.ghanaweb.com/GhanaHomePage/features/artikel.php?ID=216789.

14. Lee, J., & Mannering, F.L. (2002). Impact of roadside features on the frequency and severity of run-off-roadside accidents: An empirical analysis. *Accident Analysis and Prevention,* 34 (2), 149–161.

15. Mercier, C. R., Shelley, M. C., Rimkus, J., & Mercier, J. M. (1997). Age and gender as predictors of injury severity in head-on highway vehicular collisions. The 76th Annual Meeting, Transportation Research Board, Washington, D.C.

16. Milton, J., & Mannering, F. (1997). The relationship among highway geometric, traffic-related elements and motor-vehicle accident frequencies. The 76th Annual Meeting, Transportation Research Board, Washington, D.C.

17. Motor Traffic and Transport Unit of the Ghana Police Service, Northern regional office, Tamale. Accident data, 2007 to 2009.

18. Nassar, S. A., Saccomanno, F.F., & Shortreed, J.H. (1997). Integrated risk model (ARM) of Ontario. The 76th Annual Meeting, Transportation Research Board, Washington, D.C.
19. Pawinee, I. (2003). Analysis of Motorcycle Accidents in Developing Countries: A case study of Khon Kaen, Thailand. *Journal of the Eastern Asia Society for Transportation Studies*, Vol.5, October.
20. Robert, B. (2000): Traffic fatalities and injuries - Are reductions the result of ‘improvements’ in highway design standards? Centre for transport studies dept. of civil and environmental engineering. Imperial college of science, technology and medicine. London.
21. Safecarguide (2004): Retrieved January 22, 2012 from the World Wide Web http://www. safecarguide.com/exp/intro/idx.htm.
22. Sarpong, S. (2011, March 10). Road accident on the increase. Retrieved December 21, 2011, from http://www.newsfromafrica.org/newsfromafrica/articles/art_594.html
23. Savolainen, P., & Mannering, F. (2007). Probabilistic models of motorcyclists’ injury severities in single and multi-vehicle crashes. *Accident Analysis and Prevention*, 39, 955-963.
24. Soyoung J., Xiao Q., & David A. N. (2009). Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis and Prevention*, 42 (2010) 213–224.
25. Tesema,T. B., Abraham, A., & Grosan, C. (2005). Rule mining and classification of road traffic accidents using adaptive regression trees. *International Journal of Simulation Systems*. Science & Technology Special Issue on Soft Computing for Modeling and Simulation, 6, 10-11.

26. Tewolde M. (2007). Empirical analysis on traffic accidents involving human injuries. The case of Addis Ababa. University of Addis Ababa, Ethiopia.

27. The national roads authority: High accident locations. Retrieved November 22, 2011, from the World Wide Web www.nra.ie/PublicationsResources/DownloadableDocumentation/RoadSafety/file.

28. Vamsi K. I. (2010). *Identification of factors related to motorcycle fatal injuries* in Ohio, University Of Dayton. Dayton, Ohio.

29. Williams, M., & Hoffmann, E. (1978). Motorcycle conspicuity and traffic accidents. *Accident Analysis and Prevention*, 11, 209-224.

30. World Health Organization. (2009). *Global status report on road safety*. Retrieved November 7, 2011, from the World Wide Web http://www.who.int/violence_injury_prevention/road_safety_status/2009/en/

31. World Health Organization. (2011). *World report on road traffic injury prevention*. Retrieved January 22, 2012, from the World Wide Web http://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en/index.html

32. Yannis, G., Golias, J., & Papadimitri, E. (2005). Driver age and vehicle engine size effects on fault and severity in young motorcyclists accidents. *Accident Analysis and Prevention*, 35, 327-333.

33. Yau, K. (2003). Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. *Accident Analysis and Prevention*, 36, 333-340.

APPENDIX A

FATALITY FORM OF THE UNIT

The following information is what is required of any accident involved driver;

1. Sex of driver

2. Age of driver—————

3. Name of driver—————

4. Date of accident—————

5. Use of alcohol when accident occurred. a) Yes b) No

6. Use of safety belt when accident occurred. a) Yes b) No

7. Vehicle ownership—————

8. Type of vehicle—————

9. Age of vehicle —————

10. Weight of vehicle—————

11. Vehicles tyres condition

a) New or good as new b) Somewhat worn d) Quite worn e) Very worn

12. Estimated Speed of vehicle when accident occurred

13. Posted speed limit at the site of accident

14. Estimated annual vehicle kilometers driven—————

15. Consequences of the accident

a) Number of minor injuries—————b) Number of serious injuries ————— c) Number of fatalities —————

16. Road surface condition when accident occurred

17. Weather condition when accident occurred

18. Traffic lighting condition when accident occurred

19. Drivers familiarity of route. a) Pass scene of accident at least once a month. b) more seldom pass than once in a month

20. Type of driving license—————

21. Age of driving license—————

22. Scene of accident. a) Junction b) link c) other

23. cause of accident

a) Excessive speeding

b) Inattention, confusion of lack of judgment of driver

c) Drivers careless at road junction and cutting corners

d) Improperly overtaking or cutting in

e) Inexperience of driver

f) Intoxication

g) Other recklessness or negligence by drivers

h) Over loading

i) Mechanical defects

j) Defective lights

k) Dazzling lights

l) Skid and road surface defects

m) Other road defects

n) Obstructions

o) Level erosions

p) Children

q) Adults crossing road carelessly

r) Adults boarding or alighting from vehicles

s) Other pedestrian faults

t) Passengers faults

u) Animals not under control

v) Recklessness or negligence caused by peddle cyclist

w) Recklessness or negligence caused by drivers of horse - driven vehicles

APPENDIX B

SPSS OUTPUT OF RESULTS

[DataSet2] C:\Users\user\Documents\Data.sav

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	398	100.0
	Missing Cases	0	.0
	Total	398	100.0
Unselected Cases		0	.0
Total		398	100.0

a. If weight is in effect, see classification table for the total number of cases.

b. The variable Lerosion is constant for the selected cases. Since a constant term was specified, the variable will be removed from the analysis.

c. The variable AdultsBV is constant for the selected cases. Since a constant term was specified, the variable will be removed from the analysis.

d. The variable Animals is constant for the selected cases. Since a constant term was specified, the variable will be removed from the analysis.

e. The variable HorseDrivenV is constant for the selected cases. Since a constant term was specified, the variable will be removed from the analysis.

Dependent Variable Encoding

Original Value	Internal Value
No fatality	0
Fatality	1

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	4.364	.450	94.042	1	.000	78.600

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	34.207[a]	.048	.379

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender	18.736	5787.218	.000	1	.997	1.371E8
	Age	-.014	.052	.073	1	.787	.986
	EsiveSpeed	-2.395	2.110	1.289	1	.256	.091
	Inattentn	13.839	9640.361	.000	1	.999	1023837.090
	CarelessJ	14.117	7159.984	.000	1	.998	1351482.719
	WOvertaking	-2.911	1.929	2.278	1	.131	.054
	Inexperience	-3.633	1.903	3.643	1	.056	.026
	Intoxication	14.474	4116.495	.000	1	.997	1931722.343
	ORecklness	14.226	4631.994	.000	1	.998	1507534.975
	OverLoading	-6.407	2.413	7.049	1	.008	.002
	MechDefects	-3.608	2.021	3.186	1	.074	.027
	LightsDefects	17.592	40192.969	.000	1	1.000	4.365E7
	RSurfDefects	-4.696	28997.819	.000	1	1.000	.009
	ORDefects	13.694	19037.554	.000	1	.999	885773.331
	Obstructn	-4.851	2.390	4.121	1	.042	.008

Children	14.160	28373.733	.000	1	1.000	1410780.912
AdultsCRC	13.140	11290.749	.000	1	.999	508803.884
OPedestrian	14.054	40192.970	.000	1	1.000	1268914.865
PassengerF	-18.806	57135.291	.000	1	1.000	.000
PeddleCyclist	13.467	21569.898	.000	1	1.000	706022.050
Constant	7.567	2.350	10.370	1	.001	1934.147

a. Variable(s) entered on step 1: Gender, Age, EsiveSpeed, Inattentn, CarelessJ, WOvertaking, Inexperience, Intoxication, ORecklness, OverLoading, MechDefects, LightsDefects, RSurfDefects, ORDefects, Obstructn, Children, AdultsCRC, OPedestrian, PassengerF, PeddleCyclist.

Logistic Regression

[DataSet2] C:\Users\user\Documents\Data.sav

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	398	100.0
	Missing Cases	0	.0
	Total	398	100.0
Unselected Cases		0	.0
Total		398	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
No fatality	0
Fatality	1

Block 0: Beginning Block

Classification Table[a,b]

	Observed		Predicted: Coonsequences of accident: No fatality	Predicted: Coonsequences of accident: Fatality	Predicted: Percentage Correct
Step 0	Coonsequences of accident	No fatality	0	5	.0
		Fatality	0	393	100.0
	Overall Percentage				98.7

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	4.364	.450	94.042	1	.000	78.600

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	OverLoading	25.073	1	.000
		Obstructn	4.996	1	.025
	Overall Statistics		30.419	2	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	7.931	2	.019
	Block	7.931	2	.019
	Model	7.931	2	.019

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	45.776[a]	.020	.156

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	0	.

Contingency Table for Hosmer and Lemeshow Test

		Coonsequences of accident = No fatality		Coonsequences of accident = Fatality		
		Observed	Expected	Observed	Expected	Total
Step 1	1	2	2.000	13	13.000	15
	2	3	3.000	380	380.000	383

Classification Table[a]

	Predicted		
	Coonsequences of accident		Percentage Correct
Observed	No fatality	Fatality	

Step 1	Coonsequences of accident	No fatality	0	5	.0
		Fatality	0	393	100.0
	Overall Percentage				98.7

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	OverLoading	-4.148	1.355	9.373	1	.002	.016	.001	.225
	Obstructn	-2.444	1.195	4.185	1	.041	.087	.008	.903
	Constant	4.842	.580	69.771	1	.000	126.667		

a. Variable(s) entered on step 1: OverLoading, Obstructn.