<u>SPATIO-TEMPORAL ANALYSIS OF ROAD TRAFFIC CRASHES IN KENYA: EVIDENCE FROM GEO-REFERENCED ACCIDENT REPORT</u>

STUDENT NAME: PETER MUNYALO

REG NUMBER: AST/5457/22

DATE:12/22/2025

DECLARATION

I hereby declare that this thesis is my original work and has not been presented for the award of a degree or any other academic qualification in this or any other institution of higher learning. All sources of information used in this thesis have been duly acknowledged through appropriate referencing.

This thesis was carried out under the supervision of my supervisor(s) and complies with the regulations and requirements of the university regarding academic integrity and originality.

Name: Peter Munyalo
Registration Number: _AST/5457/22
Degree Programme: B.Sc. Applied Statistics with Computing
Institution: Moi University

Signature:
Date:

DEDICATION

This thesis is dedicated to my family for their unconditional love, encouragement, and support throughout my academic journey. Their patience, sacrifices, and belief in me gave me the strength to complete this study. I also dedicate this work to all road users, especially pedestrians and motorcyclists, whose safety and wellbeing inspired this research.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to the Almighty God for His guidance, strength, and grace throughout the entire period of my study. Without His provision of wisdom, good health, and perseverance, the successful completion of this thesis would not have been possible.

I am deeply grateful to my Dr Edwin Chirchir for his invaluable guidance, constructive criticism, and continuous support throughout the research process. His academic insight and encouragement greatly contributed to the quality of this work.

I also wish to acknowledge the support and cooperation of the university lecturers and staff whose assistance, advice, and resources were instrumental during my course of study. Special thanks go to my classmates and friends for their moral support, encouragement, and helpful discussions.

Finally, I am profoundly thankful to my family for their unwavering support, prayers, patience, and encouragement throughout my academic journey. Their belief in me provided the motivation to persevere and successfully complete this project.

Table of Contents

Table of Contents

List of Tables

List of Figures

ABSTRACT

Road traffic accidents remain a major public safety concern in Kenya, contributing significantly to fatalities, injuries, and economic losses. Despite ongoing road safety interventions, crash severity and fatal outcomes continue to persist, particularly among vulnerable road users. This study conducts a spatio-temporal and severity analysis of road traffic accidents in Kenya using geo-referenced crash data.

The study utilizes a quantitative, observational research design based on secondary data comprising 2,595 road traffic crash records obtained from crowdsourced Ma3Route accident reports. Descriptive statistics, temporal analysis, spatial hotspot mapping, correlation analysis, and logistic regression were employed to examine accident patterns and factors associated with fatal crash outcomes. Geographic coordinates were used to visualize accident hotspots, while regression techniques were applied to assess the likelihood of fatal crashes based on road user involvement, reporting intensity, and time of occurrence.

The findings reveal distinct spatial hotspots along major transport corridors and urban areas. Pedestrian and motorcycle involvement were found to be strongly associated with fatal crash outcomes, highlighting the heightened vulnerability of these road users. Additionally, fatal crashes were associated with higher reporting intensity, suggesting that the number of crash reports may serve as a proxy indicator of crash severity. Temporal factors influenced accident frequency but showed weaker associations with fatality likelihood.

The study demonstrates the value of data-driven approaches in understanding road traffic accident patterns and provides evidence-based insights to inform targeted road safety interventions and policy formulation in Kenya.

CHAPTER ONE: INTRODUCTION

1.1 Background of the study

Road traffic accidents constitute a major public safety and development challenge globally, with low- and middle-income countries bearing a disproportionate share of fatalities and injuries. In Kenya, road traffic crashes remain one of the leading causes of death and injury, particularly among economically productive age groups. The consequences of road accidents extend beyond loss of life to include long-term injuries, psychological trauma, traffic congestion, and significant economic costs related to healthcare, property damage, and lost productivity.

Despite the implementation of road safety policies, traffic regulations, and public awareness campaigns, the incidence of severe and fatal road traffic accidents in Kenya remains high. One of the key challenges facing road safety planning is the limited use of data-driven evidence to guide targeted interventions. Many existing strategies are applied uniformly without sufficient understanding of where accidents occur most frequently, when they are most likely to happen, and which road users are most vulnerable.

Advances in data availability, geospatial technologies, and data analytics provide new opportunities to analyze road traffic accidents in greater depth. Geo-referenced crash data, combined with temporal and severity indicators, can be used to identify accident hotspots, peak risk periods, and factors associated with fatal outcomes. Crowdsourced traffic reporting platforms such as Ma3Route have further expanded access to real-time and location-specific crash information in Kenya.

This study leverages geo-referenced road traffic accident data to conduct a spatio-temporal and severity analysis of road crashes in Kenya. By applying statistical and geospatial analysis techniques, the study seeks to generate evidence-based insights that can support targeted road safety interventions and contribute to reducing road traffic fatalities.

1.2 Statement of Problem

Road traffic accidents remain a persistent and critical public safety problem in Kenya, contributing to high levels of fatalities, injuries, and economic losses. Despite the implementation of traffic regulations, infrastructure improvements, and road safety awareness campaigns, the frequency and severity of road traffic crashes continue to be high, particularly along major highways and in urban areas. Vulnerable road users such as pedestrians and motorcyclists are disproportionately affected, yet targeted interventions addressing their specific risks remain limited.

One of the key challenges in addressing road traffic accidents in Kenya is the lack of comprehensive, data-driven evidence to inform policy and planning. Many road safety interventions are implemented without adequate understanding of the spatial distribution of accidents, temporal risk patterns, and factors associated with fatal crash outcomes. Existing studies often rely on aggregated statistics that fail to capture localized hotspots, peak accident periods, and road user-specific risk factors.

The growing availability of geo-referenced and crowdsourced crash data presents an opportunity to bridge this gap. However, such data remains underutilized in systematic spatio-temporal and severity analyses. Without empirical insights derived from detailed data analysis, road safety policies may remain inefficient and poorly targeted.

This study seeks to address this problem by analyzing geo-referenced road traffic accident data in Kenya to identify spatial hotspots, temporal patterns, and factors associated with fatal crash outcomes, thereby providing evidence-based insights to support more effective road safety interventions and policy formulation.

1.3 Research Objectives

1.3.1 General Objectives

To analyze the spatial, temporal, and severity patterns of road traffic accidents in Kenya using geo-referenced crash data in order to inform evidence-based road safety interventions.

1.3.2 Specific Objectives

- To examine the temporal distribution of road traffic accidents in Kenya by hour, day, and month.
- To identify spatial hotspots of road traffic accidents across different locations in Kenya using geographic coordinates.
- To assess the relationship between different road user categories (pedestrians, motorcyclists, matatus) and crash fatality outcomes.
- To determine the factors associated with the likelihood of fatal road traffic accidents using correlation and logistic regression analysis.
- To provide data-driven insights and policy recommendations aimed at improving road safety and reducing traffic-related fatalities in Kenya.

1.4 Research Hypotheses

The study was guided by the following hypotheses, which were tested at a 5% level of significance ($\alpha = 0.05$):

H_{01}: There is no significant relationship between the time of occurrence and the frequency of road traffic accidents in Kenya.

H_{11}: There is a significant relationship between the time of occurrence and the frequency of road traffic accidents in Kenya.

H_{02}: There is no significant association between pedestrian involvement and fatal road traffic accidents in Kenya.

H_{12}: There is a significant association between pedestrian involvement and fatal road traffic accidents in Kenya.

H_{03}: There is no significant association between motorcycle involvement and fatal road traffic accidents in Kenya.

H_{13}: There is a significant association between motorcycle involvement and fatal road traffic accidents in Kenya.

H_{04}: There is no significant relationship between the number of crash reports and fatal road traffic accident outcomes.

H_{14}: There is a significant relationship between the number of crash reports and fatal road traffic accident outcomes.

H_{05}: The selected explanatory variables do not significantly predict the likelihood of fatal road traffic accidents in Kenya.

H_{15}: The selected explanatory variables significantly predict the likelihood of fatal road traffic accidents in Kenya

1.5 Significance of the Study

The findings of this study are expected to be valuable to several stakeholders involved in road safety planning, policy formulation, and research in Kenya. By providing empirical evidence on the spatial, temporal, and severity patterns of road traffic accidents, the study contributes to a better understanding of factors associated with fatal crash outcomes.

For government agencies and policymakers, including the National Transport and Safety Authority (NTSA) and county governments, the results will provide data-driven insights that can support the design of targeted road safety interventions. Identification of accident hotspots and high-risk time periods can inform infrastructure improvements, traffic enforcement strategies, and allocation of limited road safety resources.

For road safety practitioners and planners, the study highlights vulnerable road users such as pedestrians and motorcyclists, enabling the development of focused awareness campaigns and protective measures aimed at reducing fatalities among these groups.

For researchers and academics, the study contributes to existing literature by demonstrating the application of spatio-temporal analysis, correlation analysis, and logistic regression techniques in road safety research using geo-referenced and crowdsourced data. It also provides a methodological reference for similar studies in developing country contexts.

For students and data science practitioners, the study serves as a practical example of applying data analytics and geospatial techniques to address real-world societal problems. Overall, the study supports evidence-based decision-making and contributes to ongoing efforts aimed at reducing road traffic accidents and improving road safety in Kenya.

1.6 Scope of the Study

This study focuses on the analysis of road traffic accidents in Kenya using geo-referenced crash data obtained from crowdsourced *Ma3Route accident reports*. The scope of the study is limited to examining the spatial, temporal, and severity characteristics of reported road traffic accidents across different locations in Kenya.

The analysis covers accident occurrences based on geographic coordinates, time of occurrence, road user involvement (including pedestrians, motorcyclists, and matatus), reporting intensity, and crash fatality outcomes. Statistical and data analytics techniques, including descriptive analysis, correlation analysis, spatial hotspot mapping, and logistic regression, are applied to identify patterns and factors associated with fatal road traffic accidents.

The study is limited to secondary data and does not involve primary data collection, field surveys, or interviews. Additionally, the analysis relies on reported crash information and does not include detailed clinical injury severity data or official police accident records. The findings are therefore interpreted within the context of the available dataset and its limitations.

Geographically, the study covers reported road traffic accidents across Kenya without focusing on a specific county or region. Temporally, the analysis is restricted to the period covered by the available dataset. The study is conducted for academic purposes and aims to generate insights that can support road safety research, planning, and policy formulation in Kenya.

1.7 Limitation of the Study

Despite providing valuable insights, this study has several limitations that should be considered when interpreting the findings. First, the study relies on secondary, crowdsourced data obtained from Ma3Route accident reports. Such data may be subject to underreporting, reporting bias, or inconsistencies, as not all road traffic accidents are reported, and reporting accuracy depends on users.

Second, the dataset does not include official police or hospital records, which limits the ability to verify accident details such as confirmed fatalities, injury severity, or post-crash outcomes. As a result, crash fatality classification is based on reported indicators rather than medically verified data.

Third, the geographic coordinates used in the analysis depend on the accuracy of user-reported locations. Some accident locations may therefore be imprecise, which could affect the identification of exact accident hotspots.

Fourth, the study does not incorporate road infrastructure characteristics such as road condition, signage, lighting, or traffic volume, which are known to influence accident occurrence and severity.

Fifth, the analysis is limited to the time period covered by the available dataset, and therefore may not fully capture long-term trends in road traffic accidents in Kenya.

Finally, the study focuses on selected variables and does not account for driver behavior factors such as speeding, alcohol use, or driver experience, which may also influence crash severity.

Despite these limitations, the study provides meaningful data-driven insights that contribute to understanding road traffic accident patterns and support evidence-based road safety interventions in Kenya.

1.8 Organization of the Study

This study is organized into five chapters, each focusing on a specific aspect of the research.

Chapter One presents the introduction to the study. It provides the background of the study, statement of the problem, research objectives, research hypotheses, significance of the study, scope of the study, and limitations of the study. This chapter sets the foundation and context for the research.

Chapter Two reviews relevant literature related to road traffic accidents and road safety. It discusses theoretical frameworks, empirical studies on spatial and temporal analysis of road traffic accidents, factors influencing crash severity, and identifies research gaps that the current study seeks to address.

Chapter Three outlines the research methodology adopted in the study. It describes the research design, data sources, study variables, data collection procedures, data analysis techniques, and ethical considerations.

Chapter Four presents the results and findings of the study. This chapter includes descriptive statistics, temporal and spatial analysis results, correlation analysis, logistic regression results, and interpretation of the findings in relation to the research objectives and hypotheses.

Chapter Five provides a summary of the study, conclusions drawn from the findings, and policy recommendations aimed at improving road safety in Kenya. It also suggests areas for further research.

CHAPTER TWO: LITERATURE VIEW

2.1 Introduction

This chapter presents a review of relevant literature related to road traffic accidents, with a specific focus on spatial, temporal, and severity analysis of road traffic crashes. The purpose of the literature review is to examine existing theoretical and empirical studies in order to establish the current state of knowledge, identify key factors influencing road traffic accidents, and highlight gaps that justify the present study.

Road traffic accidents have been widely studied due to their significant impact on public health, economic development, and transport systems, particularly in low- and middle-income countries. Previous research has explored various dimensions of road traffic accidents, including temporal patterns, spatial distribution, road user vulnerability, and factors influencing crash severity. Advances in data analytics and geospatial technologies have further enhanced the ability to analyze accident data and identify high-risk locations and periods.

This chapter reviews literature under key thematic areas, including the global and regional context of road traffic accidents, theoretical frameworks relevant to road safety analysis, spatio-temporal analysis of road traffic crashes, factors associated with crash severity, and the application of statistical and geospatial methods in road accident studies. Emphasis is placed on studies conducted in developing countries and Kenya, where road traffic accidents remain a major challenge.

By critically reviewing existing studies, this chapter identifies methodological approaches, key findings, and research gaps that inform the design and focus of the current study. The review therefore provides a conceptual and empirical foundation for the methodology and analysis presented in subsequent chapters.

2.2Theoretical Literature Review

Theoretical frameworks provide a foundation for understanding the causes, patterns, and severity of road traffic accidents. This study is guided by several theories that explain the interaction between road users, vehicles, infrastructure, and the environment in influencing accident occurrence and outcomes.

2..2.1 Systems Theory of Road Safety

The Systems Theory of Road Safety explains that road accidents are not caused by a single factor, but by the interaction of many elements within the road transport system. These elements include road users, vehicles, road infrastructure, traffic laws, and the surrounding environment. According to this theory, improving road safety requires addressing all parts of the system rather than focusing on only one factor, such as driver behavior.

The theory emphasizes that if one part of the system fails, it can increase the risk of accidents. Therefore, road safety interventions should be coordinated and comprehensive, involving engineering, enforcement, education, and emergency response (OECD, 1997).

2.2.2 Haddon's Matrix Theory

Haddon's Matrix Theory was developed to analyze road traffic injuries by considering different factors before, during, and after a crash. The matrix is divided into three phases (pre-crash, crash, and post-crash) and three influencing factors (human, vehicle, and environment). This creates a structured framework for identifying where and how accidents can be prevented or their impacts reduced.

The theory helps policymakers and researchers design effective road safety measures such as driver training, vehicle safety features, road design improvements, and emergency medical response. It is widely used in traffic safety studies because it provides a clear and systematic way of understanding accident causes (Haddon, 1980)

2.2.3 Human Factors Theory

Human Factors Theory focuses on how human behavior and limitations contribute to road traffic accidents. It explains that errors such as speeding, distraction, fatigue, alcohol use, and poor decision-making increase the likelihood of accidents. The theory recognizes that humans are prone to mistakes, especially under stress or poor road conditions.

According to this theory, improving road safety requires better driver education, enforcement of traffic rules, and designing roads and vehicles that reduce the chances of human error. The theory is important in understanding how driver behavior influences accident occurrence and severity (Reason, 1990).

2.2.4 Spatial Interaction Theory

Spatial Interaction Theory explains how movement between different locations affects accident occurrence. The theory suggests that traffic accidents are more likely to occur in areas with high interaction between people, vehicles, and activities, such as urban centers, major highways, and transport corridors.

The theory is useful for analyzing the spatial distribution of road traffic accidents and identifying accident hotspots. It helps explain why certain locations experience more accidents than others due to traffic volume, land use patterns, and connectivity between places (Ullman, 1956).

2.3 Empirical Literature Review

Empirical studies on road traffic accidents have focused on the **causes, patterns, spatial distribution, and impacts** of accidents. Researchers have used statistical, spatial, and econometric methods to understand the factors that influence accident occurrence and severity.

2.3.1 Road User Factors and Traffic Accidents

Several studies show that **human behavior** is a major contributor to road traffic accidents. According to Peden et al. (2004), risky behaviors such as speeding, alcohol consumption, reckless driving, and failure to obey traffic rules significantly increase the likelihood of accidents. Similarly, Abdel-Aty and Radwan (2000) found that driver age, experience, and reaction time strongly influence accident severity. These studies conclude that improving driver behavior through training and enforcement can reduce road accidents.

2.3.2 Vehicle and Road Infrastructure Factors

Previous research indicates that **vehicle condition and road infrastructure** play an important role in road safety. A study by Elvik, Høye, Vaa, and Sørensen (2009) revealed that poor vehicle maintenance and lack of safety features increase injury severity during crashes. In addition, road conditions such as narrow lanes, poor signage, potholes, and lack of street lighting have been linked to higher accident rates (Ackaah & Salifu, 2011). These findings highlight the importance of proper road design and vehicle inspection.

2.3.3 Environmental and Weather Factors

Environmental factors such as **rain, fog, and poor visibility** have also been associated with road traffic accidents. Qiu and Nixon (2008) found that accident frequency increases during adverse weather conditions due to reduced visibility and slippery roads. The study emphasizes the need for weather-responsive traffic management systems and driver awareness during unfavorable conditions.

2.3.4 Spatial Distribution and Accident Hotspots

Several studies have applied **spatial analysis techniques** to identify accident hotspots. Anderson (2009) used Geographic Information Systems (GIS) to map accident locations and found that accidents tend to cluster around busy intersections and major highways. Similarly, Erdogan, Yilmaz, Baybura, and Gullu (2008) showed that hotspot analysis helps identify high-risk areas for targeted interventions. These studies support the use of spatial analysis in road safety planning.

2.3.5 Statistical and Regression Analysis of Road Accidents

Empirical studies have also used **regression models** to analyze factors influencing accident occurrence. Miaou and Lum (1993) applied logistic regression models and found that traffic volume, road geometry, and speed limits significantly affect accident probability. More recent studies confirm that regression analysis is effective in predicting accident risk and supporting evidence-based policy decisions (Washington, Karlaftis, & Mannering, 2011).

2.4 Research Gaps

Despite the extensive literature on road traffic accidents, several gaps remain that justify the current study.

First, many studies focus mainly on **national or international-level data**, with limited attention to **localized or area-specific analysis**. This limits the understanding of accident patterns at the local level, where targeted interventions are most effective.

Second, although several studies examine accident causes, **few integrate statistical analysis with spatial hotspot analysis** to clearly identify high-risk locations and link them with contributing factors.

Third, most previous studies rely heavily on **descriptive analysis**, with limited use of **advanced models such as logistic regression** to examine the likelihood of accident occurrence based on multiple influencing variables.

Fourth, some studies focus on either human factors or environmental factors independently, without sufficiently examining their **combined effects** within a single analytical framework.

Finally, there is limited research evaluating how findings from accident analysis can be translated into **practical, evidence-based policy recommendations** for road safety planning and management.

2.5 Summary of literature and research gaps

2.5.1 Summary of Literature Review

The reviewed literature shows that road traffic accidents are a major public safety problem worldwide, especially in developing countries. Previous studies agree that road accidents are caused by a combination of **human factors, vehicle conditions, road and environmental factors, and traffic management systems**.

Many studies highlight **human factors** such as speeding, careless driving, alcohol use, fatigue, and non-compliance with traffic rules as major contributors to road accidents (Reason, 1990; Peden et al., 2004). Other studies emphasize the role of **road conditions**, including poor road design, lack of signage, potholes, and inadequate lighting, in increasing accident risks (OECD, 1997).

Research has also shown that **traffic volume and weather conditions** significantly influence accident occurrence, with higher accidents reported during peak traffic hours and adverse weather conditions such as rain and fog (Golob & Recker, 2003). In addition, spatial studies indicate that road accidents tend to concentrate in specific locations such as busy intersections, highways, and urban centers, forming identifiable accident hotspots (Anderson, 2009).

Methodologically, previous studies have applied **descriptive statistics, correlation analysis, regression models, and spatial analysis techniques** to examine accident patterns and contributing factors. These methods have helped improve understanding of accident trends and risk factors across different regions.

Overall, the reviewed literature provides strong evidence that road traffic accidents are a result of multiple interacting factors and that effective road safety interventions must be multi-dimensional.

2.5.2 Summary of Research gaps

Although many studies have examined road traffic accidents, gaps still exist. Most studies focus on either behavioral or infrastructural factors separately, with limited integration of **spatial and statistical models**. In addition, there is limited localized research in many developing regions. This study addresses these gaps by combining descriptive statistics, spatial hotspot analysis, correlation analysis, and logistic regression.

2.6 Conceptual Framework

A conceptual framework provides a visual and explanatory representation of the relationship between the key variables of a study. It illustrates how independent variables influence the dependent variable, while acknowledging the role of intervening factors. In this study, the conceptual framework is guided by road safety and accident causation theories and focuses on factors influencing the severity of road traffic accidents in Kenya.

Independent Variables

The independent variables in this study include:

- **Road user characteristics** (pedestrians, motorcyclists, matatus, other vehicles)
- **Temporal factors** (time of day, day of the week)
- **Spatial factors** (accident location, geographic coordinates, accident hotspots)

These variables are expected to influence the likelihood and severity of road traffic accidents.

Dependent Variable

The dependent variable is:

- **Crash severity**, measured by whether a road traffic accident resulted in a fatal or non-fatal outcome.

Intervening Variables

Intervening variables may influence the relationship between the independent and dependent variables. These include:

- Road infrastructure conditions
- Traffic enforcement and compliance with traffic laws
- Emergency response time and access to medical services

Although these variables may not be directly measured in this study, they are recognized as important contextual factors affecting crash outcomes.

Explanation of the Framework

The conceptual framework assumes that road traffic accident severity is influenced by the interaction of road user involvement, temporal patterns, and spatial characteristics. For instance, accidents involving vulnerable road users such as pedestrians and motorcyclists, occurring during high-risk periods or within identified hotspots, are more likely to result in fatal outcomes. Correlation analysis and logistic regression are used to examine the strength and direction of these relationships.

CHAPTER THREE: RESEARCH METHODOLOGY

3.1 Introduction

This chapter presents the research methodology adopted in the study to analyze road traffic accidents in Kenya. It outlines the research design, data sources, study area, data collection procedures, and analytical techniques used to achieve the study objectives. The methodology is guided by the conceptual framework and is designed to systematically examine the spatial, temporal, and severity patterns of road traffic accidents.

The chapter further describes the variables used in the analysis and explains the statistical and spatial methods applied, including descriptive analysis, correlation analysis, and logistic regression. These methods are employed to identify accident hotspots and determine factors associated with fatal road traffic accidents. The chapter also discusses ethical considerations and limitations related to data use.

Overall, this methodology provides a structured and rigorous approach for generating reliable, evidence-based findings that support informed road safety interventions and policy formulation in Kenya.

3.2 Research Design

This study adopts a **quantitative research design** with a **descriptive and explanatory approach** to analyze road traffic accidents in Kenya. The quantitative design is appropriate because the study relies on numerical and geo-referenced secondary data to examine patterns, relationships, and factors associated with road traffic accident severity.

The descriptive component of the research design is used to analyze and summarize the spatial and temporal characteristics of road traffic accidents. This includes identifying accident hotspots, examining time-based trends, and describing the distribution of accidents by road user type. Descriptive statistics and spatial analysis techniques are applied to provide an overall understanding of accident patterns.

The explanatory component of the design is employed to investigate the relationships between selected independent variables and accident severity. Correlation analysis is used to assess the strength and direction of relationships among variables, while logistic regression analysis is applied to determine factors associated with the likelihood of fatal road traffic accidents. This approach enables the study to test hypotheses and draw inferences about the determinants of crash severity.

The research design is cross-sectional in nature, as the analysis is based on accident data collected over a specified period. Overall, this design is suitable for achieving the study objectives and providing evidence-based insights to inform road safety interventions and policy decisions in Kenya.

3.3 Target Population

The target population for this study comprises **all reported road traffic accidents in Kenya** within the period covered by the dataset used in the analysis. The population includes accidents involving different categories of road users such as pedestrians, motorcyclists, matatu occupants, and other motorists, across various geographic locations in the country.

Since the study relies on **secondary, geo-referenced crash data**, the target population is defined in terms of **road traffic accident records** rather than individual persons. Each record represents a reported road traffic crash, including information on location, time of occurrence, road user involvement, and accident severity.

This target population is appropriate for the study because it captures a wide range of road traffic accidents occurring in both urban and rural areas of Kenya. Analyzing this population enables the study to examine spatial and temporal patterns of accidents and to identify factors associated with fatal and non-fatal crash outcomes.

3.4 Data source and Description

This study utilizes secondary data obtained from a geo-referenced road traffic crash database covering reported road traffic accidents in Kenya during the study period. The dataset contains systematically recorded information on road traffic accidents and is suitable for analyzing spatial, temporal, and severity patterns.

The data include key variables such as geographic location (latitude and longitude), time and date of occurrence, road user type involved, and crash severity (fatal or non-fatal). These variables are essential for identifying accident hotspots, examining temporal trends, and assessing factors associated with fatal road traffic accidents.

Prior to analysis, the data were cleaned and processed to ensure accuracy and consistency. Records with missing or incomplete key information, such as location or crash severity, were excluded from the analysis. The cleaned dataset was then prepared for spatial analysis, correlation analysis, and logistic regression.

The use of secondary, geo-referenced crash data enables comprehensive coverage of road traffic accidents across different regions of Kenya. This data source provides a reliable basis for evidence-based analysis and supports the study’s objective of informing targeted road safety interventions and policy formulation.

3.5Reliability and Validity of Data Collection Instrument

Reliability and validity are essential in ensuring the credibility and accuracy of research findings. In this study, the data collection instrument consisted of a **secondary, geo-referenced road traffic accident database**, which is routinely compiled and maintained by relevant authorities.

Reliability

Refers to the consistency and dependability of the data. The dataset used in this study is considered reliable because it is compiled using standardized reporting procedures for road traffic accidents. Accident records are consistently captured using uniform formats that include location, time of occurrence, road user involvement, and crash severity. Additionally, data cleaning procedures were undertaken to remove incomplete or inconsistent records, further enhancing the reliability of the dataset for analysis

Validity

Refers to the extent to which the data accurately measure what they are intended to measure. The dataset demonstrates content validity as it contains variables that directly align with the study objectives, including crash severity, temporal factors, spatial location, and road user type. Construct validity is ensured by clearly defining and coding variables such as fatal and non-fatal accidents based on established criteria. The use of geo-referenced coordinates enhances validity by enabling precise spatial analysis and hotspot identification.

3.6 Data Analysis and Presentation

Data analysis in this study involved both **descriptive and inferential statistical techniques**, supported by spatial analysis, to address the research objectives and test the stated hypotheses. The analysis was conducted using appropriate statistical and GIS tools to examine spatial, temporal, and severity patterns of road traffic accidents in Kenya.

Descriptive analysis was used to summarize and present the characteristics of road traffic accidents. This included frequency distributions, percentages, and graphical presentations to describe accident occurrence by location, time of day, road user category, and crash severity. Spatial analysis techniques were applied to map accident locations and identify high-risk areas (hotspots) using geo-referenced data.

To examine relationships among variables, **correlation analysis** was conducted to determine the strength and direction of associations between selected independent variables and accident severity. This provided preliminary insights into potential risk factors associated with fatal road traffic accidents.

For inferential analysis, **logistic regression analysis** was employed to assess factors influencing the likelihood of fatal road traffic accidents. Accident severity was treated as a binary dependent variable (fatal or non-fatal), while selected temporal, spatial, and road user characteristics were included as independent variables. The regression results were used to estimate odds ratios and identify statistically significant predictors of fatal crashes.

The findings of the analysis were presented using **tables, charts, graphs, and maps** to enhance clarity and interpretation. Statistical results were reported at appropriate significance levels, and spatial outputs were presented in map form to support policy-relevant conclusions.

3.7 Ethical Considerations

This study adhered to established ethical standards to ensure integrity, confidentiality, and responsible use of data throughout the research process. Ethical approval was sought from the relevant academic authority prior to data collection and analysis.

The study relied on **secondary data** obtained from officially recognized institutions responsible for road traffic accident records. Permission to access and use the data was obtained where required, and the data were used strictly for academic purposes. No primary data involving direct interaction with human participants was collected.

Confidentiality and anonymity were maintained by ensuring that the data did not contain personally identifiable information such as names, vehicle registration numbers, or contact details of accident victims. The analysis focused solely on aggregated and anonymized records to prevent identification of individuals or specific households.

Data integrity was ensured by accurate handling, processing, and analysis of the datasets. The researcher avoided data manipulation, misrepresentation, or selective reporting of findings. All sources of data and related literature were properly acknowledged to avoid plagiarism.

Finally, the study findings were reported objectively and responsibly, with due consideration of their implications for road safety policy and public welfare. The research aimed to contribute positively to knowledge and support evidence-based decision-making without causing harm to individuals or institutions involved.

CHAPTER FOUR: DATA ANALYSIS, INTERPRETATION AND DISCUSSION

4.1 Introduction

This chapter presents the analysis, interpretation, and discussion of the findings of the study based on the objectives and hypotheses outlined in the earlier chapters. The analysis focuses on examining the spatial, temporal, and severity patterns of road traffic accidents in Kenya using the geo-referenced crash data described in Chapter Three.

The chapter begins with descriptive analysis to summarize the characteristics of road traffic accidents, including their distribution by location, time of occurrence, road user category, and crash severity. This is followed by inferential analysis, including correlation analysis and logistic regression, to assess relationships between selected independent variables and the likelihood of fatal road traffic accidents.

The results are presented using tables, charts, and maps to enhance clarity and understanding. Interpretation of the findings is done in relation to the study objectives, hypotheses, and existing theoretical and empirical literature. The discussion highlights key patterns, explains observed trends, and compares the findings with previous studies.

Overall, this chapter provides empirical evidence on factors influencing road traffic accident severity in Kenya and forms the basis for conclusions and recommendations presented in the subsequent chapter.

4.2 Descriptive Analysis of Road Traffic Accidents

This section presents the descriptive analysis of road traffic accidents based on the geo-referenced crash data used in the study. The analysis provides an overview of accident characteristics and highlights patterns related to crash severity, road user involvement, time of occurrence, and spatial distribution. Descriptive statistics are used to summarize the data using frequencies, percentages, tables, charts, and maps.

4.2.1 Distribution of Road Traffic Accidents by Crash Severity

The analysis shows that road traffic accidents in Kenya comprise both fatal and non-fatal outcomes. Non-fatal accidents account for the majority of reported cases, while fatal accidents constitute a smaller but significant proportion. This indicates that although most crashes do not result in death, fatal accidents remain a major road safety concern due to their social and economic impacts.

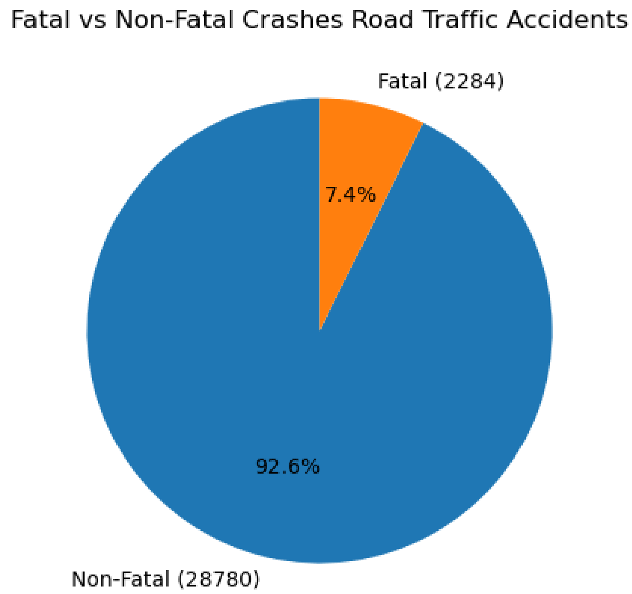

Figure 1 shows the proportion of fatal and non-fatal road traffic accidents. The results indicate that non-fatal accidents constitute a larger share of the total accidents, while fatal accidents represent a smaller proportion.

Figure 1

4.2.2 Distribution by Road User Category

Road traffic accidents involved different categories of road users, including pedestrians, motorcyclists, matatu occupants, and other motorists. The results indicate that **vulnerable road users**, particularly pedestrians and motorcyclists, were involved in a substantial proportion of accidents. These groups also exhibited a higher likelihood of fatal outcomes compared to occupants of other vehicles, highlighting their increased exposure and vulnerability on Kenyan roads.

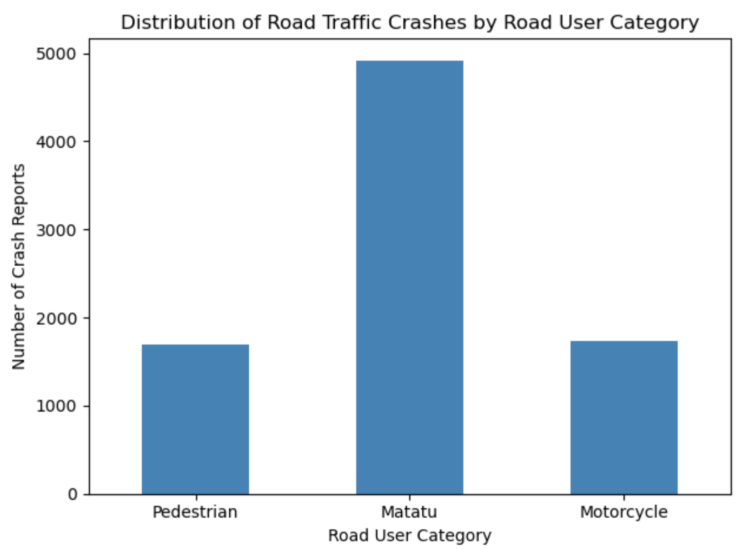

Figure 2 illustrates the distribution of road traffic crashes by road user category. The findings indicate that Matau-related crashes account for a substantial proportion of reported accidents, followed by Motorcycle-related crashes, while Pedestrian-related crashes constitute a smaller share of the total crash reports.

Figure 2

4.2.3 Temporal Distribution of Road Traffic Accidents

Analysis of the temporal characteristics of road traffic accidents revealed variations by **time of day and day of the week**. A higher number of accidents were observed during peak traffic periods, particularly in the morning and evening hours. Accidents were also more frequent on weekdays compared to weekends, reflecting increased traffic volume associated with work and school travel.

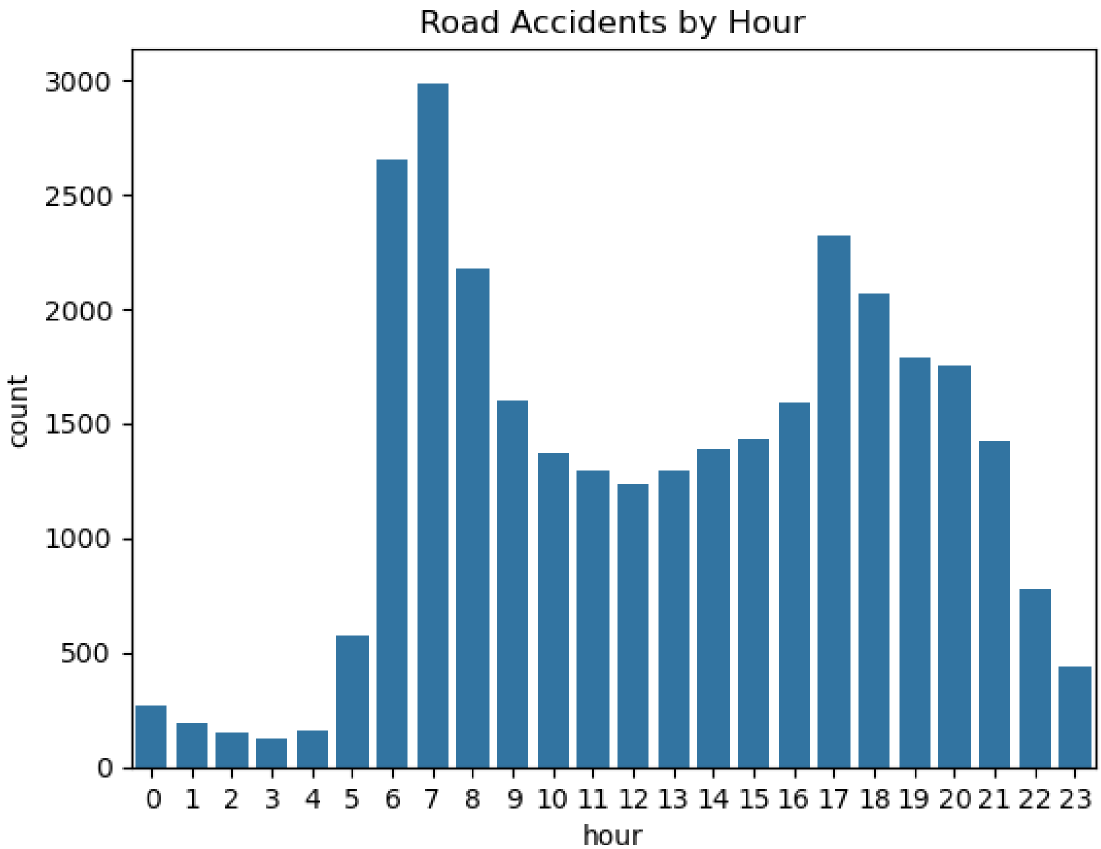

Figure 3

Figure 3 shows accident occurrences are relatively **low during the late night and early morning hours** (midnight to about 4:00 a.m.), which may be attributed to reduced traffic volume during these hours. A sharp increase in accidents is observed during the **early morning hours between 6:00 a.m. and 8:00 a.m.**, corresponding to the morning rush hour when traffic volume is high and road users are commuting to work and school.

After this peak, accident counts decline slightly during the late morning and early afternoon but rise again in the **late afternoon and early evening hours (approximately 4:00 p.m. to 7:00 p.m.)**. This second peak coincides with the evening rush hour, increased traffic congestion, driver fatigue, and reduced visibility as daylight fades. Accident frequencies then gradually decrease towards late evening hours.

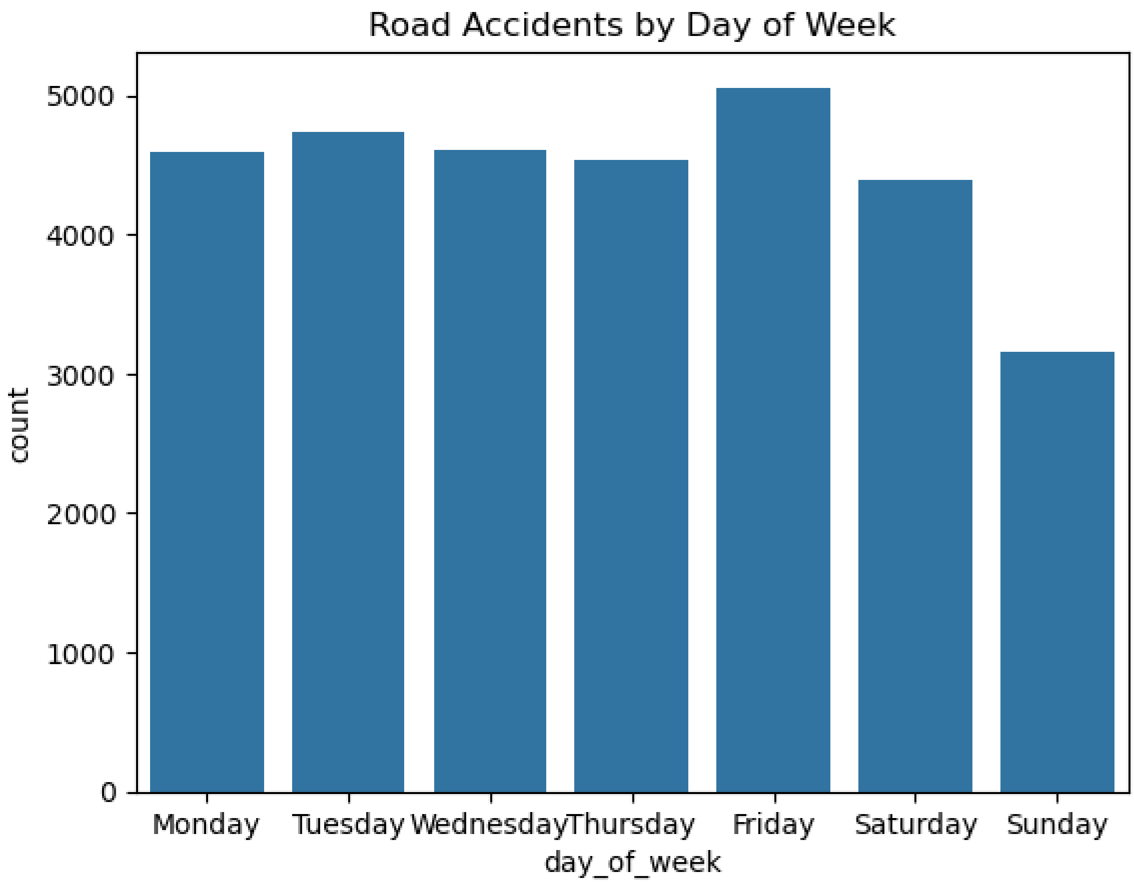

Figure 4

Figure 4 shows the distribution of road traffic accidents by day of the week. The findings show that accidents occur more frequently on **weekdays than on weekends**.

Accident counts are relatively high from **Monday to Friday**, with the highest number recorded on **Friday**. This pattern may be linked to increased traffic flow during weekdays due to work-related travel and social activities toward the end of the workweek.

A noticeable decline in accidents is observed on **Saturday**, with the **lowest number recorded on Sunday**. The reduced accident frequency on weekends, particularly Sundays, may be attributed to lower traffic volumes and fewer commercial and commuter activities.

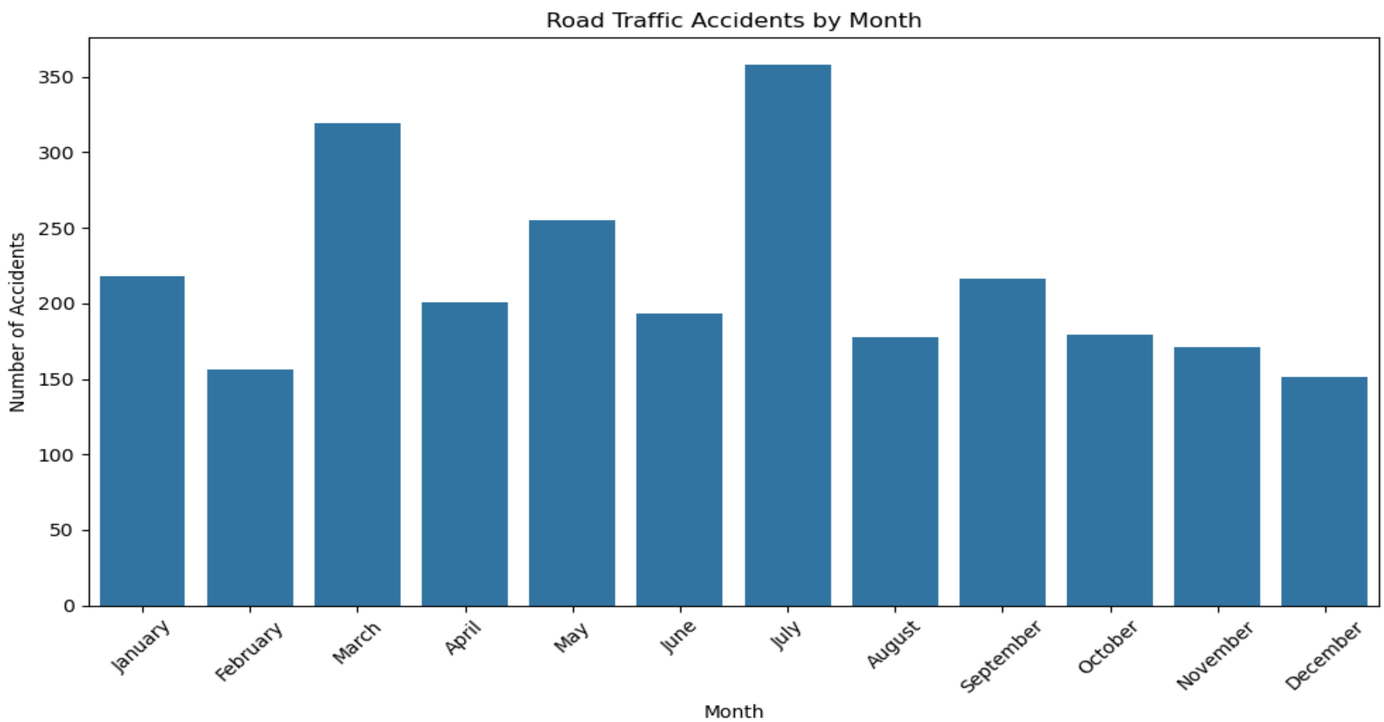

Figure 5

Figure 5illustrates the monthly distribution of reported road traffic accidents . Accident occurrences vary across the year, with noticeable peaks around March and July, and lower counts towards the end of the year, particularly in December. These patterns may be associated with seasonal factors such as increased travel during school terms, long-distance transport activities, weather conditions, and variations in traffic volume. Overall, the figure highlights clear temporal (monthly) variation in road traffic accidents in Kenya, underscoring the importance of time-based road safety interventions.

4.2.4 Severity & Risk Factor Analysis

This section looks at how serious road accidents are in Kenya and what leads to deaths. Using location-based crash data, it focuses on how different road users affect the chance of a fatal accident. The analysis mainly examines pedestrians, motorcyclists and matatus by comparing fatal and non-fatal crashes. This helps identify which road users face the highest risk and supports targeted road safety measures to reduce deaths on Kenyan roads.

Fatality vs motorcyclist

Table 1

Motorcyclist involve	NO	YES
Fatality		
NO	94.12	80.0
YES	5.88	20.0

Table 1 shows the relationship between motorcycle involvement and fatality occurrence in road traffic crashes in Kenya. Most non-fatal crashes (94.12%) do not involve motorcycles, while motorcycle involvement is relatively higher in fatal crashes, accounting for 20% of fatality-related reports. This suggests that crashes involving motorcycles are associated with an increased risk of fatal outcomes, highlighting motorcyclists as a vulnerable road user group in Kenya.

Fatality vs Pedestrian

Table 2

Pedestrian involve	NO	YES
Fatality		
NO	94.277822	65.57377
YES	5.722178	34.42623

Table 2 illustrates the relationship between pedestrian involvement and fatality occurrence in road traffic crashes in Kenya. While most non-fatal crashes do not involve pedestrians (94.28%), pedestrian involvement is substantially higher in fatal crashes, accounting for 34.43% of fatality-related reports. This indicates that crashes involving pedestrians are strongly associated with fatal outcomes, underscoring pedestrians as one of the most vulnerable road user groups in Kenya.

Fatality vs Matatu

Table 3

Matatu involve	NO	YES
Fatality		
NO	93.629505	93.301435
YES	6.370495	6.698565

Table 3 shows the relationship between matatu involvement and fatality occurrence in road traffic crashes in Kenya. Matatus appear in a similar proportion of both non-fatal (93.63%) and fatal (93.30%) crash reports, with only a slight increase in fatal crashes (6.70%). This suggests that, unlike pedestrians and motorcyclists, matatu involvement is not strongly associated with a higher likelihood of fatal outcomes in the reported crashes.

4.2.5 Spatial Distribution of Road Traffic Accidents

The spatial analysis indicates that road traffic accidents are not evenly distributed across the study area. Higher concentrations of accidents were observed along major highways, urban centers, and high-traffic corridors. These areas were identified as accident hotspots, suggesting increased risk due to high traffic volumes, mixed road user activity, and infrastructural challenges.

4.2.6 Summary of Descriptive Findings

The descriptive analysis reveals clear patterns in the occurrence and severity of road traffic accidents in Kenya. Overall, non-fatal accidents constitute the majority of reported crashes, while fatal accidents account for a relatively smaller proportion. In terms of road user categories, matatu-related crashes form the largest share of reported accidents, followed by motorcycle-related crashes, with pedestrian-related crashes making up a smaller proportion of total incidents.

Temporal analysis shows distinct daily, weekly, and monthly variations. Accident frequencies are lowest during late night and early morning hours, increase sharply during the morning rush hour, decline slightly during mid-day, and peak again during the evening rush hour, reflecting periods of high traffic volume, congestion, and fatigue. Weekday accidents are more frequent than weekend accidents, with the highest numbers recorded on Fridays, while Sundays record the fewest crashes. Monthly patterns indicate higher accident occurrences around March and

July, with lower counts toward the end of the year, suggesting the influence of seasonal travel patterns and traffic volume.

Severity analysis highlights that crashes involving motorcycles and pedestrians are disproportionately associated with fatal outcomes, underscoring these groups as particularly vulnerable road users in Kenya. In contrast, matatu involvement does not show a strong association with fatality occurrence. Spatial analysis further demonstrates that road traffic accidents are unevenly distributed, with hotspots concentrated along major highways, urban areas, and high-traffic corridors. These findings emphasize the combined influence of road user vulnerability, temporal factors, and spatial characteristics on road traffic crash risk and severity in Kenya

4.2.7 Spatial Distribution and Accident Hotspot Analysis

This section presents the spatial analysis of road traffic accidents based on the geo-referenced crash data used in the study. Spatial analysis was conducted to examine the geographic distribution of accidents and to identify high-risk locations (hotspots) where road traffic accidents are concentrated.

4.2.7.1 Spatial Distribution of Road Traffic Accidents

The spatial distribution of road traffic accidents along the MA3 route in Kenya reveals a non-random pattern, with crashes clustering around specific segments of the corridor. Accident locations were georeferenced using latitude and longitude coordinates and visualized using Geographic Information Systems (GIS) techniques. The distribution shows a higher concentration of accidents near urban centers, major junctions, and sections characterized by high traffic volume and frequent interactions between public service vehicles, pedestrians, and private motorists.

Rural segments exhibit relatively fewer crashes; however, accidents in these areas tend to be more severe, likely due to higher travel speeds and delayed emergency response. The spatial pattern suggests that infrastructure design, traffic density, and human behavior play a critical role in influencing crash occurrence along the corridor.

For interactive map click on this link:accident_hotspot_map.html

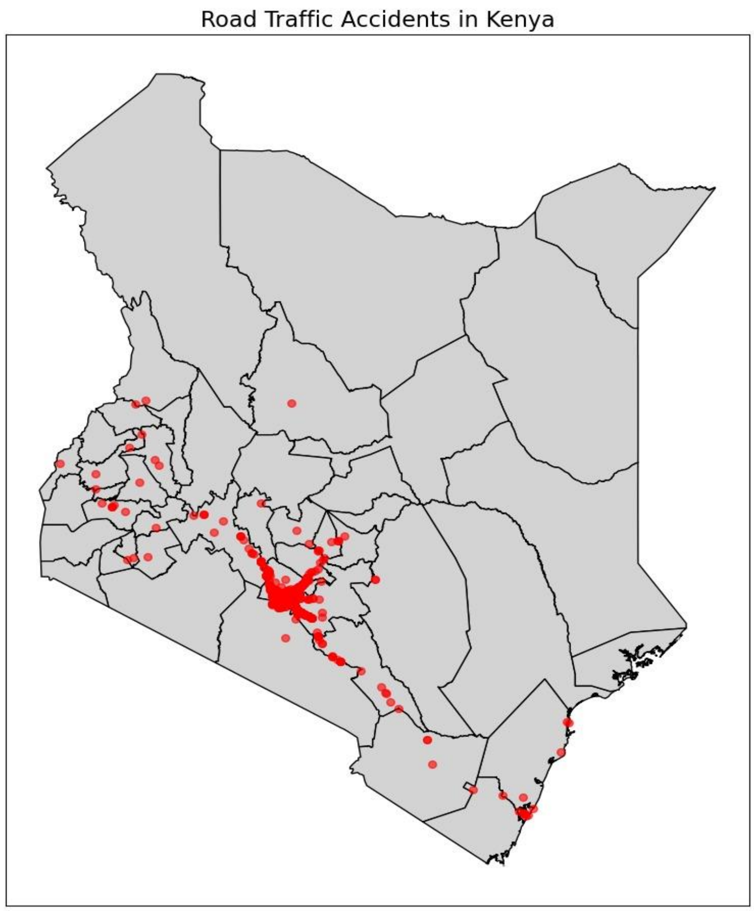

Figure 6

4.2.7.2 Accident Hotspot Analysis

From the map generated (Accident_hotspot_map), several key observations can be made:

Concentration of Accidents in Urban Areas

Most accidents appear clustered around major cities and towns. For example, areas around Nairobi, Mombasa, and Kisumu show higher densities of points, indicating that urban centers experience more road traffic incidents. This is likely due to higher traffic volumes, more vehicles, and increased interactions between different road users.

Accident Hotspots

Hotspots are regions where accidents are frequent and clustered closely together. On the map, these are shown as areas with dense overlapping markers or more intense color in the heatmap layer. Counties such as Nairobi, Kiambu, Machakos, and Mombasa emerge as primary hotspots.

Secondary hotspots appear along major highways connecting cities, including the Mombasa-Nairobi highway and the Nairobi-Eldoret route.

Rural and Low-Density Areas

Northern and northeastern regions of Kenya, such as Turkana and Marsabit, show fewer accidents. The sparse distribution in these areas can be attributed to lower population density, fewer vehicles, and less traffic congestion.

Patterns and Road User Implications

The spatial analysis indicates that accidents are not randomly distributed but are concentrated in areas with high traffic activity. Pedestrians, public transport vehicles, and private cars are most likely involved in accidents in urban hotspots, highlighting vulnerable points where road safety interventions, such as better traffic management, pedestrian crossings, and speed control, are needed.

Highway and Intersection Risk

The map also shows that major highways and intersections act as secondary hotspots. Accidents in these areas often occur due to higher speeds, increased vehicle density, and complex road designs.

Usefulness for Policy and Planning

By identifying the counties and specific locations where accidents are concentrated, policymakers and traffic authorities can prioritize interventions, such as improving road infrastructure, enforcing traffic laws, and conducting public safety campaigns in high-risk areas.

4.2.7.3Summary of Spatial Analysis Findings

The spatial distribution map and hotspot analysis clearly show that road traffic accidents in Kenya are unevenly distributed, with urban areas and major highways experiencing the highest number of incidents. Hotspot analysis helps pinpoint specific counties and regions for targeted road safety measures, making this analysis crucial for reducing fatalities and injuries.

4.2.8 Correlation analysis

The correlation analysis examines how different factors in road traffic crashes in Kenya are related. The variables analyzed include the presence of fatalities, involvement of pedestrians, motorcycles, and matatus, the number of crash reports, and the hour of the crash.

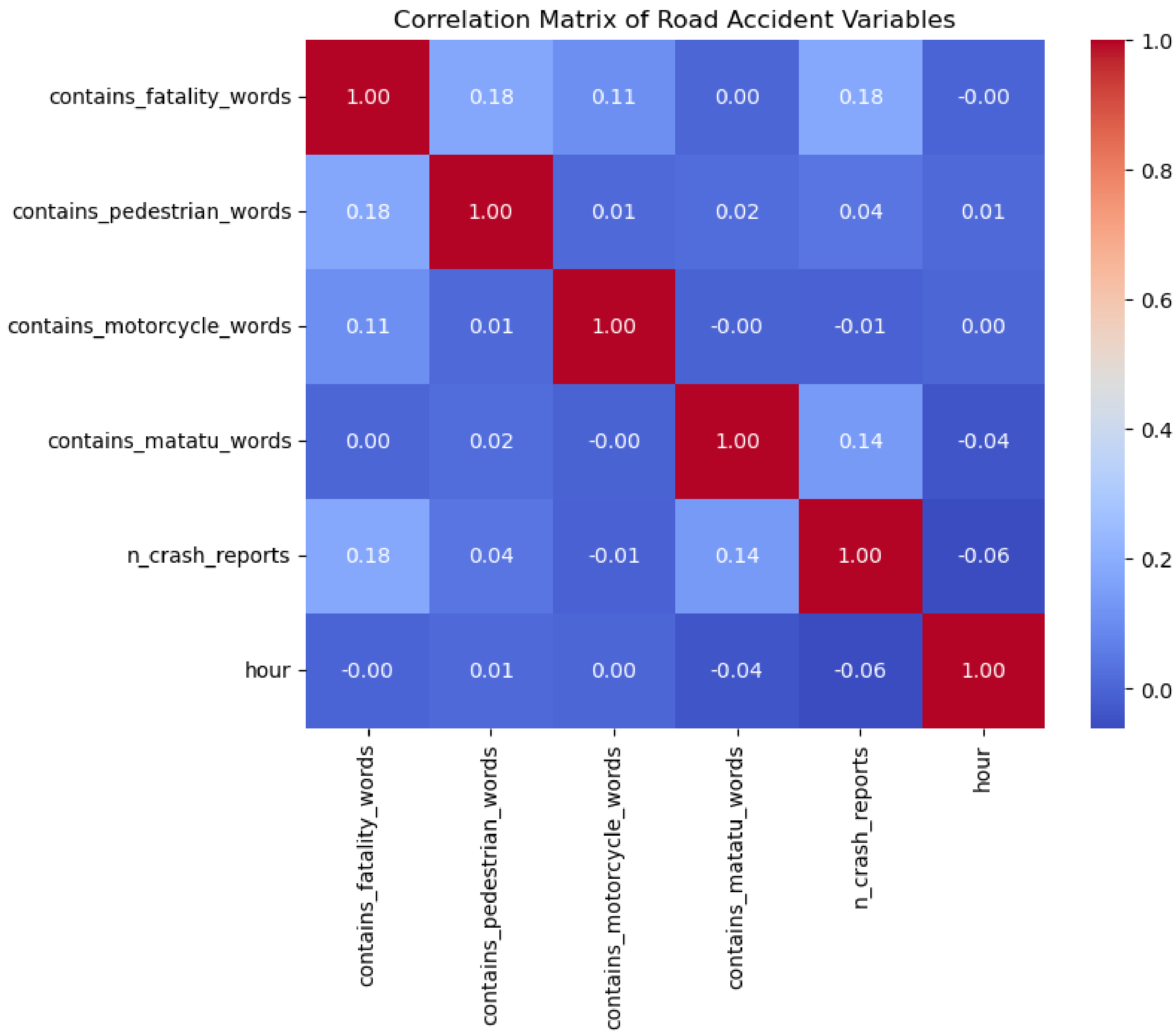

Figure 7

4.2.8.1Interpretation of correlation result

The results show that most correlations are very weak, indicating that these factors are largely independent of each other. Specifically:

i. Fatalities and other variables

Crashes that mention fatalities have a small positive correlation with pedestrian involvement (0.178) and with the total number of crash reports (0.183). This suggests that accidents involving pedestrians are slightly more likely to result in fatalities. It also indicates that areas with more crash reports tend to record slightly more fatal accidents. However, fatalities show almost no correlation with motorcycle or matatu involvement, meaning that, in this dataset, crashes involving these vehicles are not strongly linked to fatal outcomes.

ii. Pedestrian-related crashes

Pedestrian involvement shows very weak correlations with all other variables (all below 0.04). This implies that accidents involving pedestrians do not strongly coincide with other crash characteristics such as motorcycles, matatus, or the time of the crash.

iii. Motorcycle-related crashes

Motorcycle involvement also has correlations near zero with all other factors. This indicates that motorcycle crashes occur largely independently of fatalities, pedestrian or matatu involvement, or the total number of crash reports.

iv. Matatu-related crashes

Crashes involving matatus have a weak positive correlation with the number of crash reports (0.139) and a very small negative correlation with the hour of the crash (-0.036). This suggests that matatu-related crashes tend to occur slightly more in areas with many reported crashes and marginally less at later hours of the day, although the effect is very small.

v. Number of crash reports and time of crash

The total number of crash reports is weakly positively correlated with fatalities and matatu involvement, but slightly negatively correlated with the hour of the crash (-0.060). This shows that areas with more reported crashes tend to experience slightly more fatal accidents, and that crashes slightly decrease as the day progresses. The hour of the crash itself has almost no relationship with any of the crash characteristics, indicating that the time of occurrence does not strongly influence fatalities or the types of vehicles involved in crashes in this dataset.

4.2.8.2 Summary of Correlation Analysis

Overall, the analysis reveals that road traffic accidents in Kenya are mostly independent in terms of type, severity, and time. Minor patterns suggest that pedestrian involvement and high-crash areas are slightly associated with fatalities, and matatu-related crashes are marginally linked to areas with higher crash frequencies. These insights can help guide targeted interventions, such as pedestrian safety programs and focused monitoring of high-crash areas.

4.2.9 Logistic Regression Results and Discussion

A binary logistic regression model was specified with fatal crash occurrence as the dependent variable. The outcome variable indicated the presence of fatality-related terms in crash reports. Independent variables included road user type indicators, spatial coordinates, and temporal characteristics. A constant term was included to capture baseline fatality risk. The model estimates the log-odds of fatal crash occurrence as a function of these predictors.

Hypotheses Formulated

General Model Hypotheses

Null Hypothesis (H_0)

Road user type, spatial location, and temporal factors have **no significant effect** on the probability that a reported road crash is fatal.

Alternative Hypothesis (H_1)

Road user type, spatial location, and temporal factors **significantly influence** the probability that a reported road crash results in a fatal outcome.

Variable-Specific Hypotheses (Recommended)

H_{01} / H_{11}: Pedestrian Involvement

- **H_{01}:** Pedestrian involvement has **no effect** on the odds of a fatal crash.
- **H_{11}:** Pedestrian involvement **increases** the odds of a fatal crash.

H_{02} / H_{12}: Motorcyclist Involvement

- **H_{02}:** Motorcyclist involvement has **no effect** on the odds of a fatal crash.
- **H_{12}:** Motorcyclist involvement **increases** the odds of a fatal crash.

H_{03} / H_{13}: Matatu Involvement

- **H_{03}:** Matatu involvement has **no effect** on the odds of a fatal crash.
- **H_{13}:** Matatu involvement **affects** the odds of a fatal crash.

H_{04} / H_{14}: Spatial Location (Latitude & Longitude)

- **H_{04}:** Crash location (latitude and longitude) has **no effect** on fatal crash probability.
- **H_{14}:** Crash location **significantly influences** fatal crash probability.

H_{05} / H_{15}: Temporal Factors (Day & Month)

- **H_{05}:** Day of the week and month of occurrence have **no effect** on fatal crash risk.
- **H_{15}:** Temporal factors **significantly affect** fatal crash risk.

Table 4 logistic results

Variable	Odds Ratio (OR)	95% Confidence Interval	Significance	Interpretation
Pedestrian	8.72	[4.97, 15.32]	**Significant**	Crashes involving pedestrians have **about 9 times higher odds of being fatal** compared to crashes not involving pedestrians, holding other factors constant. This reflects high vulnerability due to lack of physical protection.
Motorcyclist	4.00	[2.32, 6.90]	**Significant**	Motorcyclist-involved crashes have **approximately 4 times higher odds of fatality**, indicating elevated risk associated with motorcycle use and limited protective barriers.
Matatu	0.99	[0.55, 1.79]	Not significant	Although matatus are involved in many crashes, their involvement **does not significantly affect the likelihood of a fatal outcome per crash**, suggesting high crash frequency but lower severity.
Latitude	0.62	[0.20, 1.89]	Not significant	North–south location differences do **not independently influence fatal crash risk** after controlling for other factors.
Longitude	0.50	[0.20, 1.25]	Not significant	East–west spatial variation shows **no statistically significant association** with fatal crash occurrence.
Day of Week	1.07	[0.99, 1.17]	Not significant	The day of occurrence has **no meaningful effect** on the probability of a crash being fatal.
Month	0.97	[0.92, 1.02]	Not significant	Fatal crash risk does **not vary significantly across months**, suggesting limited seasonal effects in this dataset.

Variable	Odds Ratio (OR)	95% Confidence Interval	Significance	Interpretation
Constant	4.7×10^9	Very wide CI	Not interpreted	The intercept represents baseline log-odds and has **no direct practical interpretation** in this context.

N/B Although matatus account for a large proportion of reported crashes, their involvement is not significantly associated with fatal outcomes once other factors are controlled. This suggests that matatu crashes are frequent but comparatively less severe, likely reflecting urban operating environments and higher reporting rates for non-fatal incidents.

4.2.9.1 Discussion of Logistic Regression Results

This study employed a binary logistic regression model to examine factors associated with the likelihood of a reported road crash resulting in a fatal outcome. The dependent variable captured whether a crash report contained fatality-related terms, while explanatory variables included road user type indicators, spatial coordinates, and temporal characteristics. The findings provide important insights into the determinants of crash severity, particularly highlighting the role of road user vulnerability.

Influence of Road User Type on Fatal Crash Risk

The results indicate that road user type is a critical determinant of fatal crash outcomes. Crashes involving pedestrians were found to have significantly higher odds of fatality, with an estimated odds ratio of approximately 8.7. This suggests that, holding other factors constant, pedestrian-involved crashes are nearly nine times more likely to result in a fatal outcome compared to crashes not involving pedestrians. This finding is consistent with existing road safety literature, which identifies pedestrians as highly vulnerable due to the absence of physical protection and limited capacity to withstand impact forces.

Similarly, crashes involving motorcyclists were associated with a substantially elevated fatality risk. The odds of a fatal outcome were approximately four times higher for motorcycle-related crashes compared to other crashes. This result reflects the inherent vulnerability of motorcyclists, who lack the structural protection afforded to occupants of enclosed vehicles and are often exposed to high-speed impacts. Together, these findings underscore the disproportionate risk borne by vulnerable road users and reinforce the need for targeted safety interventions aimed at pedestrians and motorcyclists.

In contrast, matatu involvement was not found to be a statistically significant predictor of fatal crash outcomes, despite matatus accounting for a large share of reported crashes. This apparent paradox can be explained by the distinction between crash frequency and crash severity. While matatus are frequently involved in reported incidents—likely due to high exposure and operation in dense urban environments—these crashes are often non-fatal. Once other factors are controlled for, matatu involvement does not significantly alter the probability that a crash results in a fatality. This finding suggests that matatu crashes are more reflective of exposure and reporting intensity rather than intrinsic fatality risk per crash.

Spatial Factors and Fatal Crash Occurrence

Spatial variables, represented by latitude and longitude, did not exhibit statistically significant associations with fatal crash risk. This indicates that, after controlling for road user type and temporal factors, broad geographic location alone does not independently explain variations in fatal outcomes. One possible explanation is that fatal crash risk is more strongly influenced by localized road conditions, traffic behavior, and enforcement levels than by general spatial positioning. The lack of significance may also reflect limited spatial variability within the study area or the need for more granular spatial indicators such as road type, speed limits, or proximity to high-risk intersections.

Temporal Characteristics of Fatal Crashes

Temporal variables, including day of the week and month of occurrence, were not statistically significant predictors of fatal crash outcomes. This suggests that fatal crash risk does not vary systematically across days or months once other factors are accounted for. While temporal patterns are often observed in crash frequency studies, their influence on crash severity appears to be weaker in this dataset. This finding implies that fatal outcomes are driven more by situational vulnerability and crash dynamics than by broad temporal trends.

Implications for Road Safety Policy and Practice

The results of the logistic regression analysis have important policy implications. The strong association between fatal outcomes and pedestrian and motorcyclist involvement highlights the urgent need for interventions targeting vulnerable road users. Such measures may include improved pedestrian infrastructure, enforcement of motorcycle helmet use, speed management in high pedestrian traffic areas, and public awareness campaigns focused on protecting non-motorized and two-wheeled road users. Conversely, the non-significant effect of matatu involvement suggests that policies aimed solely at reducing matatu crash frequency may not directly translate into reductions in fatality risk unless they also address crash severity mechanisms.

4.2.9.2 Summary of logistic regression findings

Overall, the logistic regression analysis demonstrates that fatal crash risk is primarily driven by road user vulnerability rather than spatial or temporal characteristics. Pedestrians and motorcyclists face significantly higher odds of fatal outcomes, while matatu involvement, location, and timing do not independently predict fatality risk. These findings reinforce the importance of severity-focused analysis and provide empirical support for prioritizing vulnerable road users in road safety planning and intervention strategies.

4.2.10. Negative Binomial model

Negative Binomial (NB) regression is a count data modeling technique used when the dependent variable represents non-negative integers, such as the number of road traffic accidents. It is an extension of Poisson regression that allows for **overdispersion**, a common feature in real-world crash data where the variance exceeds the mean.

In road traffic accident analysis, crash counts often vary substantially across locations and time periods due to unobserved heterogeneity, making the strict Poisson assumption of equal mean and variance unrealistic. Negative Binomial regression addresses this limitation by introducing an additional dispersion parameter that captures unexplained variability in the data.

Table 5

Variable	Coefficient	p-value	Interpretation
Intercept (const)	0.5461	0.000	Baseline crash count when all predictors = 0.
Pedestrians	0.3484	0.026	Crashes mentioning pedestrians are associated with a **41.7% increase** in crash counts.
Matatu	0.5827	0.000	Crashes involving matatus are associated with a **79% increase** in crash counts.
motorcycle	-0.0500	0.713	No statistically meaningful effect.
Hour	-0.0145	0.001	Every increase in hour slightly **reduces crash counts by 1.4%**.
Day of week	-0.0213	0.110	Not statistically significant.
Month	0.0147	0.064	Weak/ borderline effect; may show slight seasonality.

4.2.10.1 Summary of Negative Binomial model

This study employed a Negative Binomial (NB) regression model to examine factors associated with the frequency of reported road traffic accidents in Kenya. The NB model was selected due to evidence of overdispersion in the count outcome variable (*n_crash_reports*), where the variance exceeded the mean, violating the key assumption of the Poisson regression model. Compared to the Poisson model, the NB model provided a more appropriate fit by accounting for unobserved heterogeneity in crash counts.

The results indicate that crashes involving **pedestrians** and **matatus** are significantly associated with higher numbers of reported crashes. Specifically, the presence of pedestrian-related terms was associated with an increase in expected crash counts, suggesting that pedestrian involvement remains a major contributor to road traffic incidents. Similarly, matatu-related crashes showed a strong positive association, reflecting the high exposure, operational intensity, and risk profile of public transport vehicles in Kenya. These findings are consistent with existing literature that identifies pedestrians and public service vehicles as vulnerable and high-risk road user groups.

In contrast, the variable representing **motorcycle involvement** was not statistically significant in the NB model. While motorcycles are often associated with road safety concerns, this result may suggest that, after accounting for overdispersion and other covariates, motorcycle-related crashes do not significantly explain variations in the number of reported crashes in this dataset. This may also reflect reporting differences or the aggregation level of the data.

Temporal factors also played an important role. The negative coefficient for **hour of the day** indicates that crash counts tend to decrease as the day progresses, which may be linked to reduced traffic volumes during late hours or fewer crash reports at night. The **month** variable showed a marginally positive effect, suggesting seasonal variations in crash frequency, potentially driven by weather conditions, travel patterns, or holiday periods. The effect of **day of the week**, however, was not statistically significant in the NB model, indicating limited variation in crash counts across weekdays once overdispersion is accounted for.

Overall, although some variables lost statistical significance compared to the Poisson model, the NB regression provides more reliable and conservative estimates by correcting for overdispersion. The findings highlight the importance of targeting pedestrian safety and regulating public transport operations as key policy priorities. The use of the NB model strengthens the robustness of the study’s conclusions and supports evidence-based interventions aimed at reducing the frequency of road traffic accidents in Kenya.

4.2.11 Discussion of Findings

This study examined the spatio-temporal patterns and severity of road traffic crashes in Kenya using geo-referenced accident reports. The findings reveal that road traffic crashes are influenced by a combination of road user vulnerability, time, and location, with severity being driven mainly by the type of road user involved rather than when or where the crash occurs.

Overall Crash Patterns

The descriptive analysis shows that non-fatal crashes make up the majority of reported road traffic accidents in Kenya, while fatal crashes account for a smaller proportion. This pattern is expected, as many crashes result in injuries or vehicle damage without loss of life. However, the presence of fatal crashes remains a major public safety concern, especially given their social and economic impact.

Matatu-related crashes account for the highest number of reported accidents, followed by motorcycle-related crashes, while pedestrian crashes form a smaller share of total incidents. The high number of matatu crashes reflects their widespread use as a dominant mode of public transport and their frequent exposure on busy roads. However, high frequency does not necessarily translate to high severity, as shown later in the regression analysis.

Temporal Patterns of Crashes

Clear temporal trends were observed across hours of the day, days of the week, and months of the year. Accident occurrences are lowest during late night and early morning hours when traffic volumes are minimal. Crashes increase sharply during the morning rush hour, decline slightly during mid-day, and peak again during the evening rush hour. These patterns reflect periods of increased traffic congestion, time pressure, fatigue, and risky driving behavior.

Weekday crashes occur more frequently than weekend crashes, with Fridays recording the highest number of accidents and Sundays the lowest. This trend may be linked to increased travel, work-related fatigue, and end-of-week activities. Monthly patterns show higher crash frequencies around March and July, with lower counts toward the end of the year, suggesting the influence of school calendars, holiday travel, and seasonal mobility.

Severity and Vulnerable Road Users

Severity analysis highlights that pedestrians and motorcyclists are disproportionately involved in fatal crashes. Despite contributing fewer total crashes than matatus, these road users face a much higher risk of death when a crash occurs. This finding reflects their physical vulnerability, limited protection, and greater exposure to fast-moving traffic.

In contrast, matatu involvement does not show a strong association with fatal outcomes. Although matatus are involved in many crashes, most of these incidents are non-fatal. This may be due to vehicle size, structural protection, lower speeds in urban areas, or increased enforcement and regulation of public transport vehicles.

Spatial Distribution of Crashes

Spatial analysis reveals that road traffic accidents are not evenly distributed across Kenya. Crash hotspots are concentrated in urban areas, along major highways, and within high-traffic corridors. These locations experience heavy traffic flow, mixed road users, frequent intersections, and increased human activity, all of which raise the likelihood of crashes.

The hotspot analysis is particularly useful for identifying priority areas for targeted road safety interventions, such as improved road design, enforcement, pedestrian facilities, and traffic management.

Correlation Analysis Insights

The correlation analysis indicates that most crash characteristics—such as road user type, severity, and time—are largely independent of each other. However, weak positive relationships suggest that pedestrian involvement and high-crash locations are slightly associated with fatal outcomes. Matatu-related crashes also show a marginal association with areas that experience high crash frequencies. These findings suggest that while patterns exist, fatal crashes are not driven by a single factor alone.

Logistic Regression Findings

The logistic regression results provide stronger evidence on factors influencing fatal crash outcomes. The analysis shows that pedestrians and motorcyclists have significantly higher odds of being involved in fatal crashes compared to other road users. This confirms that road user vulnerability is the strongest predictor of crash severity.

Matatu involvement, spatial variables (latitude and longitude), and temporal variables (day of the week and month) do not significantly predict fatality occurrence once road user type is accounted for. This explains why matatus, despite having the highest number of reported crashes, are not statistically associated with fatal outcomes. Their crashes are frequent but generally less severe.

Negative Binomial Findings

The Negative Binomial results show that crashes involving pedestrians and matatus are significantly associated with higher reported crash counts, underscoring the vulnerability of pedestrians and the high-risk nature of public transport operations in Kenya. In contrast, motorcycle involvement was not statistically significant after accounting for overdispersion and other factors, which may reflect reporting differences or data aggregation effects rather than a lack of risk. Temporal patterns indicate that crash counts tend to decrease as the day progresses, while monthly variations suggest possible seasonal influences, whereas differences across days of the week are minimal. Overall, although fewer variables are statistically significant compared to the Poisson model, the Negative Binomial model provides more reliable and conservative estimates, reinforcing the importance of prioritizing pedestrian safety and improved regulation of public transport to reduce road traffic accidents.

Overall Implications

Taken together, these findings demonstrate that road traffic crash severity in Kenya is driven primarily by who is involved rather than when or where the crash occurs. Vulnerable road users—especially pedestrians and motorcyclists—face the greatest risk of fatal injury. Spatial and temporal factors influence where and when crashes occur but play a limited role in determining whether a crash is fatal.

These results highlight the need for road safety policies that prioritize vulnerable road users through improved pedestrian infrastructure, motorcycle safety measures, enforcement of speed limits, and targeted interventions in identified crash hotspots.

CHAPTER FIVE: SUMMARY OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents a summary of the key findings of the study, draws conclusions based on the research objectives, and provides recommendations aimed at improving road safety. The summary of findings is derived from the descriptive, spatial, correlation, and logistic regression analyses presented in Chapter Four.

The conclusions are drawn in relation to the study objectives and reflect the implications of the findings on road traffic accident occurrence. The recommendations are provided for policymakers, road safety agencies, and other stakeholders, with the aim of reducing road traffic accidents through improved traffic management, infrastructure development, and enforcement of traffic regulations. The chapter also suggests areas for further research to enhance understanding of road traffic accident dynamics.

5.2 Findings

The study revealed several important patterns regarding road traffic crashes in Kenya.

First, non-fatal crashes were found to be more common than fatal crashes. Although fatal crashes formed a smaller proportion of total accidents, they remain a major concern due to their serious social and economic consequences.

Second, analysis by road user category showed that matatu-related crashes accounted for the highest number of reported accidents, followed by motorcycle-related crashes, while pedestrian crashes made up a smaller share of total incidents. However, severity analysis showed that pedestrians and motorcyclists were more likely to be involved in fatal crashes compared to matatus. This indicates that crash frequency does not necessarily reflect crash severity.

Third, clear temporal patterns were observed. Accident occurrences were lowest during late night and early morning hours, increased during morning rush hours, and peaked again during evening rush hours. Weekday crashes were more frequent than weekend crashes, with Fridays recording the highest number of accidents and Sundays the lowest. Monthly patterns showed higher crash frequencies around March and July, suggesting the influence of seasonal travel and traffic volume.

Fourth, spatial analysis showed that road traffic accidents are unevenly distributed across Kenya. High concentrations of crashes were observed in urban areas, along major highways, and within high-traffic corridors. Hotspot analysis identified specific regions that require targeted road safety interventions.

Fifth, correlation analysis indicated that most crash characteristics were weakly related, suggesting that road traffic crashes are influenced by multiple independent factors. Slight associations were observed between pedestrian involvement and fatal outcomes, as well as between high-crash areas and matatu-related incidents.

Finally, logistic regression analysis showed that pedestrian and motorcycle involvement significantly increased the likelihood of fatal outcomes. In contrast, matatu involvement, spatial variables, and temporal variables did not significantly predict crash fatality when other factors were controlled for. This confirms that vulnerable road users face the greatest risk of fatal injury in road traffic crashes.

5.3 Conclusions

This study examined the spatio-temporal patterns, severity, and determinants of road traffic crashes in Kenya using geo-referenced secondary data. The findings confirm that road traffic accidents remain a serious public health and safety challenge, characterized by distinct spatial concentrations, temporal trends, and variations across road user groups.

The analysis revealed that non-fatal crashes constitute the majority of reported accidents, while fatal crashes, though fewer, pose significant social and economic consequences. Crash occurrence is not random; rather, it is shaped by when and where road users interact, with higher crash frequencies observed during peak traffic hours, weekdays, and in specific months, indicating the influence of traffic volume, mobility patterns, and seasonal factors.

Spatial analysis demonstrated that crashes are highly clustered in urban areas, along major highways, and within high-traffic corridors. These hotspots highlight the unequal distribution of road safety risks across the country and underscore the value of spatial analysis for identifying priority areas for targeted interventions.

Severity analysis and logistic regression results showed that vulnerable road users—particularly pedestrians and motorcyclists—face significantly higher risks of fatal outcomes compared to other road users. In contrast, matatus, despite being involved in a large number of crashes, are generally associated with non-fatal incidents. This indicates that crash severity in Kenya is driven more by road user vulnerability than by temporal or spatial factors alone.

The Negative Binomial regression further strengthened these conclusions by accounting for overdispersion in crash count data. The model showed that pedestrian- and matatu-related crashes significantly increase reported crash counts, while motorcycle involvement was not statistically significant once overdispersion and other covariates were considered. Although fewer variables were significant compared to the Poisson model, the Negative Binomial approach provided more reliable and conservative estimates, reinforcing the robustness of the study's findings.

Overall, the study concludes that improving road safety in Kenya requires a strong focus on protecting vulnerable road users, particularly pedestrians and motorcyclists, alongside targeted interventions in high-risk locations and peak traffic periods. The integration of spatial, temporal, and advanced regression analyses provides evidence-based insights that can inform effective road safety policies and contribute to reducing the burden of road traffic accidents in Kenya

5.4 Recommendations

Based on the findings and conclusions of this study, the following recommendations are proposed:

Improve Pedestrian Infrastructure

There is a need to improve pedestrian facilities such as sidewalks, footbridges, zebra crossings, and street lighting, especially in high-crash urban areas and along major highways.

Enhance Motorcycle Safety Measures

Authorities should strengthen enforcement of helmet use, promote rider training programs, and encourage the use of reflective clothing to reduce fatal injuries among motorcyclists.

Targeted Interventions in Crash Hotspots

Road safety agencies should use hotspot analysis results to prioritize high-risk locations for interventions such as speed control measures, improved road signage, and traffic calming strategies.

Strengthen Traffic Law Enforcement

Enforcement of speed limits, drunk-driving laws, and vehicle safety regulations should be intensified, particularly during peak traffic hours and on high-risk corridors.

Improve Data Collection and Reporting

Accurate and complete geo-referenced crash data should be maintained to support continuous monitoring, evaluation, and evidence-based road safety planning.

Public Awareness and Education Campaigns

Road safety education programs should focus on vulnerable road users and risky driving behaviors, targeting both drivers and pedestrians.

Strengthen Regulation and Oversight of Public Transport (Matatus)

The Negative Binomial analysis shows that matatu-related crashes significantly contribute to higher crash counts. Authorities should enhance regulation of public service vehicles through stricter enforcement of speed limits, driver training, vehicle inspections, and compliance with traffic laws. Improving operational standards in the public transport sector can substantially reduce overall crash occurrence.

Incorporate Temporal Patterns into Road Safety Planning

Temporal trends indicate higher crash occurrences during peak traffic hours and certain months of the year. Authorities should align enforcement strategies, public awareness campaigns, and traffic management efforts with these high-risk periods to improve their effectiveness.

5.5 Recommendations for further research

While this study provides important insights into the spatio-temporal patterns and severity of road traffic crashes in Kenya, several areas require further investigation to enhance understanding and inform more effective interventions.

First, future research should incorporate road environment and infrastructure variables, such as road condition, road type, presence of pedestrian facilities, signage, and lighting. Previous studies have shown that poor road design and inadequate infrastructure significantly contribute to crash occurrence and severity (WHO, 2018; Elvik et al., 2009). Including these variables would provide a more comprehensive explanation of crash risk beyond road user categories.

Second, further studies should examine driver and rider behavior factors, including speeding, alcohol and drug use, fatigue, and compliance with traffic laws. Behavioral factors have been identified as major contributors to fatal road traffic crashes, particularly in low- and middle-income countries (Peden et al., 2004; WHO, 2015). Access to police or hospital records could help capture these critical determinants of crash severity.

Third, future research could use vehicle-level data, such as vehicle age, safety features, and maintenance status, to assess how vehicle characteristics influence crash outcomes. Studies have shown that vehicle safety standards and protective features play an important role in reducing fatalities (NHTSA, 2013).

Fourth, longitudinal studies using time-series or panel data are recommended to assess trends in road traffic crashes over longer periods. Such approaches would allow researchers to evaluate the impact of policy changes, infrastructure improvements, and enforcement measures on crash frequency and severity over time (Washington et al., 2020).

Finally, future studies should explore integrated spatial modeling techniques, such as spatial regression or geographically weighted regression (GWR), to better capture spatial dependence and regional variations in crash risk. These methods can provide more precise location-based insights for targeted road safety planning (Anselin, 1988; Fotheringham et al., 2002).

Overall, expanding the scope of variables and analytical methods in future research will strengthen evidence-based decision-making and support more effective road safety strategies in Kenya.

References

Haddon, W. (1980). *Advances in the epidemiology of injuries as a basis for public policy*. Public Health Reports, 95(5), 411–421.

OECD. (1997). *Safety strategies for rural roads*. Organisation for Economic Co-operation and Development.

Reason, J. (1990). *Human error*. Cambridge University Press.

Ullman, E. L. (1956). The role of transportation and the bases for interaction. In W. L. Thomas (Ed.), *Man's role in changing the face of the Earth* (pp. 862–880). University of Chicago Press.

Abdel-Aty, M., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5), 633–642.

Ackaah, W., & Salifu, M. (2011). Crash prediction model for two-lane rural highways in Ghana. *IATSS Research*, 35(1), 34–40.

Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*, 41(3), 359–364.

Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *The handbook of road safety measures*. Emerald Group Publishing.

Erdogan, S., Yilmaz, I., Baybura, T., & Gullu, M. (2008). Geographical information systems aided traffic accident analysis system. *Accident Analysis & Prevention*, 40(1), 135–146.

Miaou, S. P., & Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention*, 25(6), 689–709.

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A. A., Jarawan, E., & Mathers, C. (2004). *World report on road traffic injury prevention*. World Health Organization.

Qiu, L., & Nixon, W. A. (2008). Effects of adverse weather on traffic crashes. *Transportation Research Record*, 2055(1), 139–146.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis*. CRC Press.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *The handbook of road safety measures* (2nd ed.). Bingley, UK: Emerald Group Publishing.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, UK: John Wiley & Sons.

National Highway Traffic Safety Administration (NHTSA). (2013). *Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey*. Washington, DC: U.S. Department of Transportation.

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A. A., Jarawan, E., & Mathers, C. (2004). *World report on road traffic injury prevention*. Geneva, Switzerland: World Health Organization.

Washington, S. P., Karlaftis, M. G., Mannering, F. L., & Anastasopoulos, P. C. (2020). *Statistical and econometric methods for transportation data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

World Health Organization. (2015). *Global status report on road safety 2015*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2018). *Global status report on road safety 2018*. Geneva, Switzerland: World Health Organization