TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

THESIS NO.: T20/079

Post Encroachment Time-Based Behavioral Analysis of Vehicle-Pedestrian Interactions at Unsignalized Midblock Crosswalks:

A Case Study of Shantinagar and Dhobighat Crosswalks in Kathmandu Valley

by

Sandesh Lamsal

A FINAL THESIS

SUBMITTED TO THE DEPARTMENT OF CIVIL ENGINEERING

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE IN TRANSPORTATION ENGINEERING

DEPARTMENT OF CIVIL ENGINEERING

LALITPUR, NEPAL

APRIL, 2025

COPYRIGHT

The author has agreed that the library, Department of Civil Engineering, Pulchowk Campus, Institute of Engineering may make this report freely available for inspection. Moreover, the author has agreed that permission for extensive copying of this thesis report for scholarly purpose may be granted by the professor(s) who supervised the thesis work recorded herein or, in their absence, by the Head of the Department wherein the thesis report was done. It is understood that the recognition will be given to the author of this report and to the Department of Civil Engineering, Pulchowk Campus, Institute of Engineering in any use of the material of this thesis report. Copying or publication or the other use of this report for financial gain without approval of the Department of Civil Engineering, Pulchowk Campus, Institute of Engineering and author's written permission is prohibited.

Request for permission to copy or to make any other use of the material in this report in whole or in part should be addressed to:

Head
Department of Civil Engineering
Pulchowk Campus, Institute of Engineering
Lalitpur, Kathmandu
Nepal

TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

DEPARTMENT OF CIVIL ENGINEERING

The undersigned certify that they have read and recommended to Institute of Engineering for acceptance, a thesis entitled **"Post Encroachment Time-Based Behavioral Analysis of Vehicle-Pedestrian Interactions at Unsignalized Midblock Crosswalks: A Case Study of Shantinagar and Dhobighat Crosswalks in Kathmandu Valley"** submitted by Sandesh Lamsal in partial fulfillment of the requirement for degree of Master of Science in Transportation Engineering.

....

Supervisor: Dr. Rojee Pradhananga
Department of Civil Engineering
Institute of Engineering

....

External Examiner: Rajesh Khadka
Program Coordinator, TEAM
Nepal Engineering College

....

[redacted] Anil Marsani
Coordinator: MSc in Transportation Engineering
Department of Civil Engineering

ABSTRACT

This study investigates pedestrian-vehicle interactions by analyzing Post-Encroachment Time (PET), a critical surrogate safety measure, across two contrasting urban environments in Kathmandu Valley: Shantinagar (urban road) and Dhobighat (arterial road). Statistical modeling and behavioral analysis identify key determinants of PET, including Curb Time, Vehicle Type, Yielding Behavior, Own Path, Conflict Type, Crossing Speed, Traffic Speed, and Perceived Volume. In Shantinagar, assertive pedestrian behavior prevails, driven by the dominance of two-wheelers, which rarely yield, prompting pedestrians to navigate traffic assertively. Conversely, Dhobighat exhibits a more cautious pedestrian profile (reserved behavior), influenced by higher traffic speeds and a larger proportion of SUV-Cars and Bus-Trucks. A multiple linear regression model explains 28.3% of PET variance, reflecting the inherently stochastic nature of pedestrian interactions. Modified Weibull Survival Analysis defines site-specific PET thresholds: 1.21 and 2.45 seconds for urban roads, and 1.71 and 2.98 seconds for arterial roads, effectively categorizing interactions into high, medium, and low-risk zones. In locations where site-specific thresholds cannot be established due to limited data, combined PET thresholds of 1.34 and 2.88 seconds derived from the overall dataset might be used. Severity analysis reveals distinct risk profiles across vehicle types. Two-wheelers are associated with predominantly low-to-moderate severity, accounting for 56% of interactions. SUV-Cars show higher risk, with 59.20% of interactions falling into the high to very high severity categories. Bus-Trucks pose the greatest danger, with 16.90% of their interactions classified as Extreme Severity. Based on these findings, we recommend adopting context-specific PET thresholds for more accurate risk classification and enforcing stricter yielding laws and speed limits near midblock crosswalks, particularly in areas with high heavy-vehicle presence.

Keywords: Vehicle-Pedestrian Interaction, Post-Encroachment Time, Conflict Dynamics, Ordinary Least Squares Regression, Survival Analysis, Severity Analysis

ACKNOWLEDGEMENT

There are many individuals to whom I am deeply indebted, without whose support this research would not have been possible. First and foremost, I would like to express my heartfelt gratitude to Dr. Rojee Pradhananga, my academic supervisor, for her invaluable knowledge and expertise. Her unwavering guidance and encouragement were instrumental in shaping this study and ensuring its successful completion.

I am also sincerely grateful to our Program Coordinator, Asst. Professor Anil Marsani, for his invaluable support during the pilot study phase of this research. His continuous guidance and insightful suggestions played a crucial role in refining my approach. Additionally, I would like to extend my appreciation to my mentors, Dr. Pradeep Kumar Shrestha, Dr. Jagat Kumar Shrestha, and Dr. Padma Bahadur Shahi, for their continuous guidance and mentorship throughout my master's degree. Their expertise and encouragement have been invaluable in my academic journey.

Name: Sandesh Lamsal
Roll no: 079MsTrE020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PET	Post Encroachment Time
PWT	Pedestrian Waiting Time
CMF	Crash-Based Modification Factor
OLS	Ordinary Least Squares
VIF	Variance Inflation Factor
MAIS	Maximum Abbreviated Injury Scale
MLR	Multiple Linear Regression
KE	Kinetic Energy
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
KDE	Kernel Density Estimation

CHAPTER 1: INTRODUCTION

1.1 Background

Road traffic injuries and fatalities remain a significant global health and development challenge, impacting millions of lives every year. According to the World Health Organization(WHO, 2023), road traffic crashes have emerged as the leading cause of death for children and young people aged 5 to 29, and they rank as the 12th leading cause of death across all age groups worldwide. Road user fatalities are distributed among different groups: 4-wheel vehicle occupants account for 30% of global road fatalities, followed by pedestrians at 23%, and powered two- and three-wheeler users at 21% (WHO, 2023). This data highlights the vulnerability of pedestrians and the critical need for road safety interventions that consider pedestrian dynamics. In India, a severe road crash happens every minute, with a fatality occurring every four minutes. Road crashes significantly impact a country's GDP, with India losing 3% of its GDP due to such incidents. Although there have been some fluctuations, both the number of crashes and injuries have shown a decreasing trend since 2010 (MORTH-2018, 2019). The study conducted by the Indian Institute of Technology (IIT) Delhi, titled “Road Safety in India: Challenges and Opportunities,” reports that in urban areas, 58% of fatalities in Mumbai and 47% in Delhi are related to pedestrian incidents occurring on the road (Mohan et al., 2015). In developing countries, the limited availability of pedestrian infrastructure—such as foot over bridges or subways—and the low preference for using them, even when available, often compel pedestrians to cross streets at grade, particularly at mid-block locations. This situation is further exacerbated by jaywalkers which increases the frequency of conflicts at these crossings and subsequently leads to a higher occurrence of severe crashes.

In Nepal, the growing rate of road traffic crashes underscores a similar crisis at the national level. Between 2001 and 2013, the Nepal Traffic Police recorded 95,902 crashes, resulting in 100,499 injuries and 14,512 fatalities. Alarmingly, the traffic death rate in Nepal escalated from 4 per 100,000 people in 2001–2002 to 7 per 100,000 in 2011–2012 (Karkee & Lee, 2016). Road Traffic Crashes have become a major concern in Kathmandu Valley, with the number of crashes rising significantly each year. Over the past decade, from FY

2070/71 to FY 2079/80, road crashes have surged by 128.48%, increasing from 4,672 to 10,675 incidents. Additionally, the number of fatalities has risen by 30.77%, from 143 deaths in FY 2070/71 to 187 in FY 2079/80 (Gautam & Joshi, 2024). These statistics reveal that motorcyclists and pedestrians are disproportionately affected, with higher incidents among men and people aged between 20 and 40. Factors such as risky pedestrian behavior, alcohol consumption, and substandard bus driving practices have been identified as potential contributors to this high crash rate. Kathmandu, Nepal's capital, faces critical road safety challenges as rapid urbanization drives population growth and transportation demand. Limited public transportation options have led many residents to rely on private vehicles, intensifying road congestion and encouraging faster driving, which raises risks for both drivers and pedestrians.

Walking is the primary mode of transportation in Kathmandu Valley. According to the Public Transport (PT) Survey conducted by JICA Study Team, 41% of trips in Kathmandu, 40% in Lalitpur, and 51% in Bhaktapur are made on foot. Despite this significant reliance on non-motorized transport (NMT), pedestrians are often overlooked in transportation planning. Similar to other developing countries, vehicle-oriented infrastructure development is prioritized, driven by its perceived economic and social benefits. As a result, NMT users face inadequate infrastructure, including poorly maintained sidewalks, limited pedestrian crossings, and a lack of dedicated pedestrian-friendly policies. The lack of investment in pedestrian safety exacerbates the risks for those who rely on walking as their primary mode of travel, particularly in high-traffic urban areas like Kathmandu(JICA, 2017). Frequent, unregulated vehicle-pedestrian interactions in these spaces, along with limited enforcement of traffic rules, create a hazardous environment where pedestrians and vehicles compete for road space. Crosswalks are essential components of the road network, designed to facilitate safe pedestrian movement amidst vehicular traffic. However, pedestrian safety is often compromised due to improper design, including the inadequate provision of guardrails and pedestrian refuge islands. These crossings can be categorized into signalized and unsignalized types. At unsignalized crosswalks, pedestrian safety is a greater concern at midblock locations than at intersections due to road geometry and higher vehicle speeds. Unlike intersections, where vehicles are more likely to slow down or stop, midblock crosswalks often expose pedestrians to faster-moving traffic, increasing the risk of severe conflicts. At unsignalized crosswalks, zebra crossings serve as a key visual cue for both pedestrians and drivers to facilitate safe road crossing. However, in Nepal, many zebra crossings are faded or poorly maintained, making them less visible to approaching vehicles. As a result, drivers often notice them only at the last moment, reducing their ability to yield or slow down in time. This lack of visibility increases the risk of pedestrian-vehicle conflicts, further compromising pedestrian safety. The severity of such conflicts is heightened when vehicles travel at higher speeds. For instance, a 10 km/h increase in the speed limit raises the likelihood of fatality by 46% (Chen et al., 2019). In 2015, approximately 48% of pedestrian crashes in Beijing occurred at midblock locations, resulting in 56% of total pedestrian fatalities.

In Nepal, traffic is characterized by a mixed flow of various vehicle types, including motorcycles, cars, buses, trucks, and non-motorized transport. Each vehicle type has distinct operational characteristics such as size, speed, maneuverability, and acceleration patterns, which contribute to unpredictable traffic conditions. This variability increases the complexity of pedestrian movement, particularly at unsignalized crossings, where pedestrians must navigate gaps between vehicles with differing speeds and stopping behaviors. The lack of lane discipline, high traffic density, and frequent encroachments further elevate the risk, making road crossing a significant safety concern for pedestrians. A key aspect of these interactions is the time gap between a pedestrian clearing a crossing point and a vehicle arriving, known as Post-Encroachment Time (PET), which serves as a critical measure of near-miss incidents. Several studies have employed Post Encroachment Time as a surrogate safety measure to assess vehicle-pedestrian conflicts (Marisamynathan & Vedagiri, 2020 ;Peesapati et al., 2018 ; Ansariyar, 2023). However, no such studies have been carried out in Nepal, where diverse traffic compositions and distinctive pedestrian behaviors highlight the need for localized research. Given these conditions, this study aims to explore the behavioral dynamics of vehicle-pedestrian interactions at unsignalized crosswalks characterized by heavy vehicular and pedestrian activity, such as those at Shantinagar and Dhobighat in the Kathmandu Valley. The insights generated from this study may contribute to urban planning and policy-making efforts aimed at enhancing pedestrian safety and optimizing the efficiency of shared road spaces.

1.2 Statement of the Problem

The interaction between vehicles and pedestrians is a major safety concern in urban areas, particularly at unsignalized crosswalks. While grade-separated crossings like Foot Over Bridges (FOBs) can reduce these interactions, their use is inconsistent. In a survey examining the frequency of Foot Over Bridge (FOB) usage, 44.99% of respondents indicated they sometimes use FOBs, 43.52% reported frequent use, only 10.27% stated they always use them, and 1.22% admitted to never using FOBs (Shilpakar & Shahi, 2022). This underscores the necessity of implementing safe, at-grade crosswalks. Some studies from India suggest that approximately 60% of pedestrian fatalities occur in urban areas, with nearly 85% of these incidents taking place at crosswalks (Bansal et al., 2018). A similar, if not more severe, situation is observed in Nepal, where pedestrian safety at crosswalks remains a significant concern. Additionally, since it is impractical to signalize every crosswalk, there is a critical need to study pedestrian-vehicle interactions at unsignalized crosswalk to ensure the safety of all road users. To address this, two unsignalized crosswalks with high levels of pedestrian-vehicle interaction—Shantinagar and Dhobighat—were selected for detailed analysis in this study.

1.3 Research Objectives

The primary aim of this research is to analyze the safety and dynamics of vehicle-pedestrian interactions at unsignalized crosswalks in Kathmandu Valley. The specific objectives are:

a. To identify and quantify the key factors influencing Post Encroachment Time in vehicle pedestrian interactions.
b. To categorize PET into different risk levels using survival analysis.
c. To perform severity analysis to detect and evaluate severe pedestrian-vehicle interactions, helping to identify high-risk scenarios.

1.4 Scope of Study

The study’s intended scope is as follows:

a. To compare vehicle and pedestrian interactions in two different urban environments inside Kathmandu Valley.
b. To understand the reasons behind riskier interactions and identify areas for interventions to enhance pedestrian safety.
c. To establish a threshold for PET tailored to the local traffic conditions in Kathmandu Valley.

1.5 Limitation of Study

The limitation of the study is:

a. The study does not account for pedestrian or driver psychological factors (e.g., risk perception, distraction, or urgency), which can significantly influence interaction dynamics.
b. The PET thresholds established are specific to the local context and may only be applicable in mixed traffic conditions similar to those observed in Kathmandu Valley, limiting their generalizability to other traffic systems.
c. The study focused on peak hours, which limited the analysis of vehicle pedestrian interactions observed during off peak hours.
d. This study focuses solely on normal pedestrians and does not account for distinctly abled individuals or cyclists.
e. The study considers conflicts only for following scenarios, as described in section 3.2

1.6 Organization of Report

This thesis report is organized into 5 chapters as follows:

Chapter 1: Introduction

This chapter provides an overview of road traffic crashes and existing pedestrian-vehicle interaction scenarios, emphasizing safety concerns at unsignalized crosswalks. It also outlines the study's objectives, scope, and limitations.

Chapter 2: Literature Review

It reviews existing studies on pedestrian and driver behavior, their interactions, conflict assessment, and severity analysis, with a focus on Post Encroachment Time.

Chapter 3: Methodology

This section details the step-by-step approach, from data collection to analysis, explaining the methods used to model pedestrian-vehicle interactions and classify conflicts based on PET.

Chapter 4: Results and Discussion

It presents and interprets the findings from data analysis, conflict categorization, and severity assessment, offering insights into pedestrian safety at unsignalized crosswalks.

Chapter 5: Conclusion and Recommendation

This chapter summarizes key findings, discusses their implications for pedestrian safety, and provides recommendations for infrastructure improvements and policy considerations.

CHAPTER 2: LITERATURE REVIEW

2.1 General

Limited research has been conducted in Nepal on vehicle-pedestrian interactions as a whole. Most studies focus primarily on pedestrian behavior, with some examining driver behavior. However, very few have explored vehicle-pedestrian interactions comprehensively, and none have specifically analyzed conflict measures such as Post Encroachment Time. To gain a deeper understanding of the issue, relevant information for the research was gathered from various existing sources, including books, academic journals, published and unpublished theses, conference and seminar papers, magazines, newspapers, online articles, and official websites. The various literatures related to pedestrian behavior, driver yielding tendencies, vehicle-pedestrian interaction dynamics, conflict measures and severity at crossings can be summarized below:

2.2 Pedestrian Behavior

Pedestrian behavior is complex and unpredictable as they stop, gets distracted and change direction unexpectedly. Pedestrian behavior is shaped by dynamic factors like vehicle speed and traffic volumes, social factors such as upbringing, education, and negligence, and environmental factors including site conditions, visibility, and road layout, all of which influence decision-making and safety at crossings. Previous studies have explored pedestrian behaviors during road crossings, often focusing on specific aspects (Bendak et al., 2021). For example, Koh et al. (2014) examined rule violations at signalized crosswalks, while Thompson et al. (2013) investigated the influence of social and technological distractions on crossing behaviors. This emphasizes how crucial it is to take a comprehensive approach that takes into account a greater range of pedestrian behaviors in order to better understand their trends and enhance crossing safety measures.

2.3 Driver Behavior

Driver's willingness to give the right of way to pedestrians at pedestrian crossing is low. Furthermore, drivers did not sufficiently lower their speeds to maintain a readiness to react to an unexpected dangerous situation. Moreover, drivers were willing to slow or stop when the speed of their vehicles was low. Drivers rarely yield to pedestrians at zebra crossings, with giving way occurring in only 5% of cases. However, when pedestrians act assertively to force vehicles to stop, they manage to pass in nearly half of such conflict situations(Várhelyi, 1998). In Nepal, key factors influencing driver yielding behavior includes the number of pedestrians crossing, pedestrian gender, age, and speed, vehicle type and speed, driver gender and age, presence of a median strip, and road marking condition (Bajracharya & Marsani, 2021).

2.4 Vehicle-Pedestrian Interaction

When pedestrians and drivers attempt to occupy the same space simultaneously, it creates a high-risk situation. Both parties may expect the other to yield, leading to hesitation, confusion, or near-collisions. This struggle for space is particularly common at unsignalized crossings, where clear priority is lacking, and can result in unsafe maneuvers, as drivers may brake abruptly while pedestrians pause or dart across. This concept aligns with Fuller's Threat Avoidance Model, which explains how drivers respond to perceived threats, including interactions with pedestrians. According to Fuller (1984), drivers continuously assess risks posed by others on the road and adjust their behavior based on threat perception and response capacity. When pedestrians and drivers attempt to occupy the same space, drivers make real-time threat assessments: if a pedestrian appears to yield, the driver may proceed without slowing down; if not, the perceived threat increases, prompting evasive actions. Fuller's model also highlights drivers' limitations in reacting to sudden pedestrian moves, as they may not fully prepare to yield until the threat becomes immediate, sometimes leaving insufficient time to stop. The interaction between pedestrian and vehicles occurs through different medium. Pedestrians often rely on eye contact (through glances or stares), hand waves, smiles, or nods. Drivers, meanwhile, communicate by flashing their lights, waving, or making eye contact (Sucha et al., 2017). However, for this communication to be effective, it must be clear and properly understood by both parties. Misinterpretations or lack of communication can lead to confusion, hesitation, or risky behavior, increasing the likelihood of crashes. In a survey, Risser (1985) revealed that 47% of the incidents occurred without communication, 11% were due to a lack of necessary communication, and 42% happened during actual communication. The Vehicle Pedestrian interaction is complex because Pedestrians and drivers consider and deal with the same situations, but from the very different perspective (Šucha, 2014). The critical role of communication in these interactions makes essential to study not only the dynamics of vehicle-pedestrian communication but also the associated conflict measures to ensure safe interactions.

2.5 Post Encroachment Time

Post Encroachment Time (PET) is the time interval between when one road user (such as a pedestrian or vehicle) exits a potential conflict point and when another road user reaches the same location. Johnsson et al. (2018) reviewed at a variety of approaches to evaluating conflicts involving vulnerable road users and came to the conclusion that no one indication is always appropriate in every situation. However, time-to-collision and post-encroachment time are among the most commonly used metrics for evaluating conflicts, regardless of the context.Arun et al. (2021) conducted a systematic review of studies from the past decade (2010–2019) focusing on conflict measures. Despite the introduction of numerous conflict measures over time, the earliest metrics, such as time-to-collision and post-encroachment time, remain the most widely used, likely due to their simplicity and ease of interpretation. Zheng et al. (2019) showed that combining time-to-collision with post-encroachment time enhances safety predictions, as each metric captures distinct dimensions of a conflict event. Time-to-collision reflects the anticipated risk during a conflict event, whereas post-encroachment time measures the actual safety margin after the vehicles have passed one another and the conflict has ended. Tageldin et al. (2017) highlighted that the choice of conflict measures depends on the traffic environment being studied. This review categorizes traffic systems into two types: organized systems, typical in countries like the United States and the United Kingdom, and less-structured systems, common in nations such as China, India, and South Korea. Since India and Nepal share similar traffic characteristics, conflict measures suitable for India are also applicable to Nepal. Time-to-collision calculations depend on precise measurements of speed and distance, which are easier in organized systems with automated tools whereas post-encroachment time requires direct observation, manageable even in less-organized environments(Arun et al., 2021). While time-to-collision and post-encroachment time are commonly used to identify various types of conflicts, most studies specifically use time-to-collision for rear-end conflicts.

Once road user trajectories intersect, post-encroachment time is often used as an additional relevant conflict measure(Mohamed & Saunier, 2018). Models that integrate PET with traffic volume characteristics exhibit superior predictive capability compared to those that rely solely on PET (Peesapati et al., 2018). Marisamynathan & Vedagiri (2020) established Threshold PET value for highly dangerous conflicts, dangerous conflicts, and no conflicts for various vehicle types using cumulative frequency distribution analysis. A PET value greater than 5.5 seconds indicates no interaction between a pedestrian and a vehicle, while a highly dangerous conflict is likely when PET is 2.0 seconds or less. The threshold PET values for highly dangerous conflicts were identified as 2.5 seconds for cars, 3.06 seconds for two-wheelers, 3.83 seconds for light commercial vehicles (LCVs), 2.37 seconds for heavy commercial vehicles (HCVs), and 2.25 seconds for autos. However, for midblock locations, there is a lack of significant research categorizing conflicts into different levels based on PET. Based on the insights from the literature review, PET along with relevant traffic metrics appears to be the most suitable conflict measure for the context of my study.

2.6 Conflict

A conflict in pedestrian-vehicle interaction refers to a situation where a pedestrian and a vehicle are moving in a way that creates a potential risk of collision, requiring one or both parties to take evasive action to avoid a crash, typically occurring at intersections or crossing points where pedestrians and vehicles share the same space; essentially, it's a near-miss situation between a pedestrian and a vehicle. Studies have demonstrated that conflict-based approaches are more effective than historical crash data. Several researchers argue that conflict-based analyses provide more accurate information for predicting pedestrian-vehicle collisions since crashes are infrequent, and analyzing crash data alone is challenging due to limited insights into pedestrian and driver behavior. However, other studies highlight the value of integrating historical crash data with conflict-based methods Additionally, researchers have compared subjective and objective assessments of traffic conflicts, finding a strong correlation between the two (Vedagiri & Kadali, 2016). Govinda et al. (2022) has classified conflicts into two categories: severe conflicts and normal conflicts. Severe conflicts occur when one or both road users must stop or adjust their speeds to avoid a collision, indicating a high likelihood of a crash. In contrast, normal conflicts occur when both the pedestrian and vehicle continue at their usual speeds, suggesting a low probability of a crash. The conflicts for my study are defined based on these assumptions.

2.7 Severity

Severity refers to the extent or seriousness of an event, condition, or impact. In the context of Road Traffic Crashes, severity indicates how serious the consequences of a crash are, typically measured in terms of injury levels, fatalities, property damage, and overall impact on safety. Various scales have been developed to classify and analyze crash severity, with KABCO and AIS (Abbreviated Injury Scale) being among the most widely used. The KABCO scale, determined by police officers at the crash scene, categorizes severity based on observed injuries and circumstances. It classifies injuries as K (Fatal), A (Incapacitating Injury), B (Non-incapacitating Injury), C (Possible Injury), and O (Property Damage Only). However, this system relies on subjective assessments, which may not always reflect the true medical severity of injuries. In contrast, the AIS scale is a medical-based injury classification developed based on threat to life and survivability rather than just incapacitation. Some injuries, such as severe internal trauma, may not be immediately visible to investigating officers but pose a significant risk to life. AIS codes are assigned by medical professionals with access to full medical records, ensuring a more accurate severity assessment (Burch et al., 2014).

2.8 Summary of Literature Review

The literature review reveals critical insights into pedestrian behavior, driver yielding tendencies, vehicle-pedestrian interaction dynamics, and conflict measures at unsignalized crossings, emphasizing their impact on safety and behavioral patterns. Pedestrian behavior is influenced by dynamic, social, and environmental factors, leading to complex and often unpredictable actions at crossings. Studies highlight the need for a comprehensive approach to better understand pedestrian trends and enhance safety measures. Driver behavior indicates a low willingness to yield to pedestrians, influenced by factors such as vehicle speed, pedestrian demographics, and road conditions. Vehicle-Pedestrian interaction is characterized by a struggle for space at unsignalized crossings, often leading to hesitation, confusion, or near-collisions. Communication between pedestrians and drivers is crucial but prone to misinterpretation, increasing crash risks. While Post-Encroachment Time has emerged as a reliable conflict metric, its application in Nepal remains underexplored. This study bridges that gap by leveraging PET to analyze interactions in mixed traffic, addressing the lack of localized data. Further, while survival analysis is well-established in medical and reliability fields, its potential to model time-dependent conflict risks (e.g., likelihood of safe crossings based on PET thresholds) in transportation is nascent. This research innovates by applying survival analysis to PET data, offering a dynamic, time-sensitive perspective on risk that transcends traditional static methods. Crucially, prior severity analyses rely solely on either conflict or crash data; this work advances the field by integrating both using conflict-derived PET values weighted by crash-based modification factors to quantify real-world injury risks This summary provides a foundation for analyzing behavioral dynamics of vehicle-pedestrian interactions at unsignalized midblock crosswalks using PET as a conflict measure, contextualized for Shantinagar and Dhobighat.

CHAPTER 3: METHODOLOGY

3.1 Research Framework

The methodology of this study is primarily based on video-graphic surveys. With the exception of crosswalk length and trap length to record vehicle speed, which were measured through a site survey, all other variables were carefully extracted from the recorded video footage. Approximately peak hours of video were recorded each day over three consecutive weekdays for 2 hours, resulting in a total of six hours of footage per study location. The adopted methodological framework is illustrated in Figure 3.1.

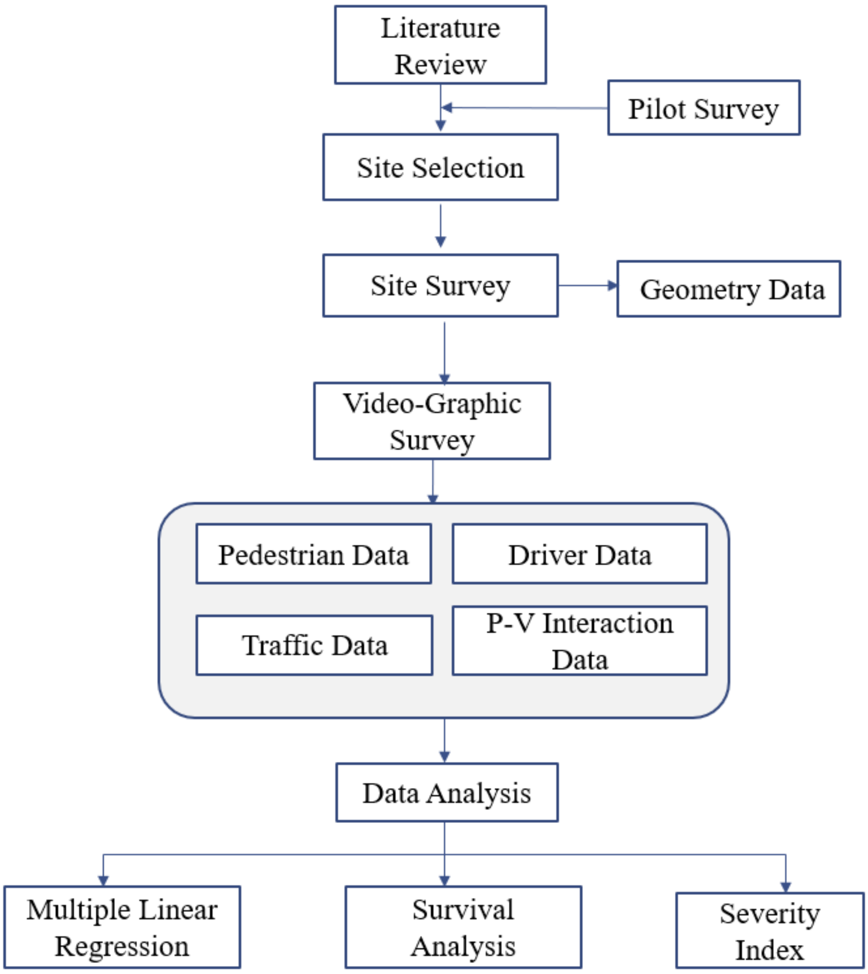


Figure 3.1 : Methodological Framework

First of all, thorough review of prior studies was conducted to identify key factors relevant to vehicle-pedestrian interactions. This review helped establish the study's framework, ensuring that all crucial variables were considered in alignment with the research objectives. The selection of study locations was based on pedestrian flow and the nature of vehicle-pedestrian interactions. Two crosswalks with distinct traffic characteristics were chosen: the Dhobighat section along the Ring Road, representing an arterial road with higher vehicular speeds, and the Shantinagar section, an urban road with lower vehicular speeds but significant pedestrian movement.

From the video footage, only the interactions that resulted in a conflict between vehicles and pedestrians were extracted, while non-conflicting interactions were not considered. The criteria for conflict in this thesis is defined in section 3.2. All relevant variables, except for the crosswalk length and vehicle trap length (which were obtained through a site survey), were carefully extracted from the footage. The extracted data was then systematically recorded on observational sheets for further analysis.

Once the necessary data was extracted, statistical analysis was performed using Python. Multiple Linear Regression was employed to examine the relationship between Post Encroachment Time and various independent variables, providing insights into the factors influencing pedestrian safety at crosswalks. In addition, survival analysis was used to assess the probability of safe crossing over time, while the severity analysis quantified the intensity of interactions. The results obtained were then interpreted and discussed in detail to highlight key findings and implications.

3.2 Variable Description

The study utilizes four categories of variables: Vehicle-Pedestrian Interaction Attribute, Pedestrian Attribute, Driver Attribute, and Traffic Attribute. These variables are used in Multiple Linear Regression, Survival Analysis and Severity Analysis . In this study, only interactions resulting in a conflict are considered. The conflict occurs under the following scenarios:

a. A vehicle approaches the crosswalk and does not yield to a pedestrian already crossing, forcing the pedestrian to stop or retreat.
b. A pedestrian steps onto the crosswalk while a vehicle is too close to stop safely.

c. A jaywalking pedestrian crosses the road, causing a vehicle to brake suddenly or maneuver.

3.2.1 Pedestrian-Vehicle Interaction Attribute

P-V (Pedestrian-Vehicle) interaction data including conflict types, conflict locations, and PET measurements examines the dynamics of encounters between pedestrians and vehicles,. These variables are crucial for assessing the severity and risk associated with these interactions.

i. Post Encroachment Time (PET)

Post Encroachment Time (PET) is the time difference between when one road user (e.g., a pedestrian) leaves the potential conflict point and when another road user (e.g., a vehicle) enters it. It serves as a key indicator of the risk of pedestrian-vehicle interactions. A PET of less than 1 second indicates high risk, 1–2 seconds is medium risk, and more than 2 seconds is low risk (VAN DER HORST, 1990). PET is considered a continuous dependent variable and is measured in seconds.

ii. Conflict Location

This variable identifies the specific lane of the road where the conflict occurs. It helps in understanding the spatial dynamics of pedestrian-vehicle interactions. It is categorized as near lane and far lane. The near lane is defined as the lane closest to where the pedestrian begins crossing, while the far lane is the lane on the opposite side of the starting point.

iii. Conflict Type

This variable categorizes the nature of conflict based on the evasive actions taken by pedestrians, drivers or both. It is categorized as maneuver conflict, speed-based conflict and hybrid conflict. Maneuver conflict occurs when one or both road users change their path to avoid a collision, whereas speed-based conflict occurs when one or both road users adjust their speed to prevent a collision. Hybrid conflict occurs when one road user changes their path while the other adjusts their speed, or when a single road user both changes their path and adjusts their speed simultaneously.

3.2.2 Pedestrian Attribute

Pedestrian data examines the characteristics and behaviors of individuals crossing the road, which influence pedestrian-vehicle interactions.

i. Gender

Gender is categorized based on the pedestrian's visual appearance as either male or female. It is an independent categorical variable.

ii. Age

Age is estimated based on visual appearance and is categorized into three groups: less than 20, 20-50, and over 50. It is an independent categorical variable.

iii. Size

Size refers to the number of pedestrians crossing simultaneously. It is classified into three categories: a single pedestrian (Size_1), two pedestrians (Size_2), and groups of more than two (Size>2).

iv. Own Path

This variable categorizes pedestrian behavior based on their sense of entitlement or confidence while crossing the road. It is categorized as assertive or reserved. Assertive refers to pedestrians who stand below the curb or use hand gestures while crossing, while reserved pedestrians stand at the curb and cross the road without using any gestures.

v. Crossing Pattern

This variable describes the manner in which pedestrians cross the road. It is categorized as either normal, where pedestrians follow road rules, or jaywalker, where they cross improperly outside designated areas and without adhering to traffic regulations.

vi. Crossing Speed

Crossing speed is the rate at which a pedestrian moves across the road, measured as a continuous independent variable in meters per second (m/s). It is calculated by dividing the total road width by the time taken to cross.

vii. Crossing Stage

Crossing stage indicates if the pedestrian crosses the road in one go or in multiple stages. It is categorized as Single or Multiple.

viii. Curb Time

Curb Time is the duration a pedestrian spends at the curb after arriving at the curb and before beginning to cross. It is an independent categorical variable categorized based on Table 3.1. It is classified as NWT (Not Waiting in Curb at all) when the pedestrian crosses without pausing at the curb, SWT (Short Waiting Time) when the pedestrian waits for 10 seconds or less at the curb, and LWT (Long Waiting Time) when the pedestrian waits for more than 10 seconds at the curb.

ix. Exposure Time

Exposure Time refers to the additional time a pedestrian spends on the road beyond the necessary crossing time, representing extra exposure to risk. In other words, it is the net time a pedestrian is actively moving across the road, excluding any waiting time in the middle of the road or below the curb. It is an independent categorical variable categorized based on Table 3.1.

Table 3.1 : Waiting Time Ranges Based on the Level of Service of Pedestrian Waiting Time (Source: Nemeth & Menk, 2014)

LOS	Descriptions	Waiting time ranges (s)
A	Usually, no conflicting traffic	0-5
B	Occasionally some delay due to conflicting traffic	5-10
C	Delay noticeable to pedestrians, but not inconveniencing	10-20
D	Delay noticeable and irritating	20-30
E	Delay approaches tolerance level, risk taking behavior likely	30-45
F	Delay exceeds tolerance level, high likelihood of pedestrian risk taking	>=45

It is classified as NET (No Extra Exposure Time) when the pedestrian crosses without any additional time spent on the road, SET (Short Extra Exposure Time) when the pedestrian spends 10 seconds or less beyond the crossing time, and LET

(Long Extra Exposure Time) when the pedestrian spends more than 10 seconds beyond the crossing time.

3.2.3 Driver Attribute

Driver attributes capture the characteristics and behaviors of drivers that influence pedestrian-vehicle interactions.

i. Traffic Speed

   Traffic Speed refers to the average speed of vehicles involved in pedestrian-vehicle interactions. It is considered a continuous independent variable, measured in meters per second (m/s). Since the focus is on the speed of vehicles directly participating in interactions, it is categorized as a driver attribute.

ii. Yield

   Yield indicates whether the interacting vehicle nearly came to a stop during the interaction, reflecting the driver's willingness to yield to a pedestrian. It is an independent categorical variable, labeled as Yes or No.

3.2.4 Traffic Attribute

Traffic Attributes capture the characteristics of the traffic environment that influence pedestrian-vehicle interactions.

i. Perceived Volume

   Perceived Volume refers to the total number of vehicles observed by a pedestrian while waiting and crossing the road in each direction. It is treated as an independent continuous variable, expressed in terms of vehicles per minute (veh/min). To capture the influence of traffic exposure on pedestrian safety, the perceived volume representing the traffic presence during the pedestrian's crossing attempt is generalized to a one-minute interval. This method is similar to the approach used by Shah & Pradhananga (2024), who calculated the number of vehicles passing in both directions while the pedestrian waited and crossed the road. However, this study introduces a modification to better reflect pedestrian behavior and exposure. Since a pedestrian's crossing decision is based on the vehicles approaching from the direction they intend to cross, the modified method distinguishes between vehicles in the near lane and the far lane. Specifically, while waiting and reaching the midpoint of the crossing, only vehicles on the near lane are counted, as the pedestrian is focused on the traffic in that lane. After reaching the midpoint and continuing the crossing, vehicles on the far lane are considered, as they are now in the pedestrian's line of sight and affect their crossing.

ii. Vehicle Type

Vehicle Type refers to the category of the vehicle involved in the interaction, identified through visual observation. It is an independent categorical variable labeled as 2W (Two-Wheeler), SUV-Car, or Bus-Truck.

3.3 Site Description

Frequent crashes at zebra crossings highlight the importance of analyzing vehicle-pedestrian interactions in urban environments. Post-Encroachment Time serves as an effective metric for evaluating the nature of these interactions. To study these dynamics in contrasting traffic scenarios within the Kathmandu Valley, two sites—Shantinagar and Dhobighat—were selected, as shown in Figure 3.2 and Figure 3.3. Both locations are situated along four-lane, undivided roads with two-way vehicular traffic, but they differ notably in traffic characteristics. Shantinagar is marked by high pedestrian volumes and dense, continuous vehicle flow, predominantly involving two-wheelers. In contrast, Dhobighat features higher vehicle speeds, greater vehicle-type diversity, and more cautious pedestrian movement, making these sites suitable for examining a wide range of interaction severities.

3.3.1 Site 1 : Shantinagar Crosswalk

Figure 3.2: Shantinagar Crosswalk

3.3.2 Site 2: Dhobighat Crosswalk

Figure 3.3: Dhobighat Crosswalk in Ring Road

3.4 Video-Graphic Survey

The video-graphic survey was conducted using a GoPro Hero 11, mounted at a suitable height to capture vehicle-pedestrian interactions clearly. A pilot survey was initially conducted for 30 minutes during both peak and off-peak hours to assess traffic conditions and evaluate the feasibility of data collection, and the impact of distant traffic signals on congestion, vehicle speed, and pedestrian movement. The findings revealed that back queuing from traffic signals had minimal impact on vehicle speed or pedestrian crossings. However, conflicts were more frequent during peak hours, consistent with previous studies that highlight higher vehicle speeds and interaction rates during this period. Based on these findings, peak hours were selected for the main analysis. After obtaining ethical approval from relevant property owners, data collection was carried out during the morning peak hours over three weekdays, with each survey session lasting for 2 hours. The morning peak hour was chosen for the main analysis as it exhibited high vehicular and pedestrian flow, making it suitable for studying interactions and conflicts. Additionally, the evening peak hour was not selected due to reduced visibility in winter, which could compromise the accuracy of video analysis. The camera was set to record in 5.3K resolution with a linear field of view at 30 fps (H.265), ensuring high-quality footage for precise data extraction and analysis.

3.5 Data Extraction

Not all interactions between pedestrians and vehicles lead to conflicts. Therefore, only those interactions that resulted in conflicts were included in the study. The scenarios that define these conflicts are outlined in Section 3.2. A total of approximately 655 data points were gathered from the Shantinagar site, while 360 data points were collected from the Dhobhighat site. These datasets were utilized for Multiple Linear Regression analysis, Survival Analysis and Severity Analysis.

3.6 Correlation Test

Before conducting multiple linear regression, it is essential to ensure that independent variables are not highly correlated, as multicollinearity can affect the reliability of the model. To assess this, a correlation test is performed. For continuous variables, Pearson's correlation coefficient is commonly used to measure the strength and direction of their linear relationship. A coefficient close to 1 indicates a strong positive correlation, while a value near -1 suggests a strong negative correlation. Ideally, the correlation should be close to 0 to minimize multicollinearity.

However, Pearson's correlation is not suitable for categorical variables. Instead, the Chi-square (χ^2) test is used to determine whether two categorical variables are statistically associated. If the test is significant, it indicates a dependency between the variables. To further assess the strength of this association, Cramér's V is used, which ranges from 0 to 1, with higher values indicating stronger relationships. A value of zero means there is no association between the variables. Values between 0.1 and 0.3 indicate a weak association, while values from 0.3 to 0.5 suggest a moderate association. If the value exceeds 0.5, it signifies a strong association between the variables

3.7 Multiple Linear Regression

Multiple linear regression is a statistical method used to examine the relationship between a continuous dependent variable and multiple independent variables. Linear regression falls under Supervised Machine Learning, where models are trained on labeled data to make predictions. Unlike classification algorithms, which predict categorical outcomes, regression algorithms focus on continuous outcomes. There are two main types of linear regression: simple linear regression, which involves one independent variable, and multiple linear regression, which includes two or more predictors.

In this study, Post Encroachment Time serves as the dependent variable, while a range of pedestrian (gender, age, size, own path, curb time , exposure time, crossing speed, crossing stage), driver (traffic speed, yield), traffic (perceived vol , vehicle type), and pedestrian-vehicle interaction (conflict type, conflict location) factors act as independent variables. This model estimates how changes in these factors influence PET by fitting the data to a linear equation. Since PET is influenced by multiple factors, this study employs multiple linear regression to quantify their combined impact on pedestrian-vehicle interactions at crosswalks. The basic equation of the MLR model is expressed in Eq. (3.1)

$$\begin{aligned} PET = \; & \beta_o + \beta_1 Gender + \beta_2 Age + \beta_3 Size + \beta_4 Own\ Path + \beta_5 Curb\ Time \\ & + \beta_6 Exposure\ Time + \beta_7 Crossing\ Speed + \beta_8 Crossing\ Stage \\ & + B_{10} Perceived\ Vol + \; \beta_{11} Traffic\ Speed + \beta_{12} Vehicle\ Type \\ & + \beta_{13} Yield + \beta_{14} Conflict\ Type + \; \epsilon \end{aligned} \tag{3.1}$$

where PET is the dependent variable

β_o is the intercept,

β_1 to β_{14} are the coefficients for each respective predictor,

ϵ is the error term

3.8 Assumption of Multiple Linear Regression

Multiple linear regression relies on several key assumptions to provide accurate predictions and valid statistical inferences. The first assumption is the linear relationship between the dependent variable (Post Encroachment Time – PET) and each independent variable, meaning that changes in the independent variables should result in proportional changes in PET. If the relationship is non-linear, scatter plots can reveal this, and transformations such as logarithmic, square root, or polynomial functions may be applied to correct it. Secondly, the independence of errors assumes that the residuals of the model are not correlated across observations. The Durbin-Watson test is commonly used to check this, with a value close to 2 indicating little to no autocorrelation. If autocorrelation is detected, lag variables, differencing the data, or time series modeling techniques can be used as corrective measures.

Another assumption is homoscedasticity, which requires the variance of residuals to remain constant across all levels of the independent variables. If heteroscedasticity is present, residual plots may show a funnel-like pattern, and solutions include log-transforming dependent variables, using weighted least squares regression, or introducing interaction terms. Additionally, the residuals of the regression model should be normally distributed, which can be checked using histograms, Q-Q plots, or the Shapiro-Wilk test. If residuals deviate from normality, outliers can be removed, or transformations and robust regression techniques can be applied.

Finally, the model assumes no perfect multicollinearity, meaning that the independent variables should not be highly correlated with each other, as this complicates the isolation of their individual effects. Variance Inflation Factor (VIF) values above 10 indicate potential multicollinearity. If detected, correlated variables should be removed, combined, or addressed using Principal Component Analysis (PCA). Ensuring that these assumptions are satisfied enhances the reliability and validity of the multiple linear regression model's predictions and inferences.

3.9 Risk Assessment Based on PET

The risk levels associated with pedestrian-vehicle interactions was assessed by categorizing PET into different risk levels. The survival analysis was employed to establish PET thresholds and categorize risk levels.

3.9.1 Weibull Survival Analysis

Weibull survival analysis is a statistical method used to model time-to-event data, where the goal is to analyze the time until an event of interest occurs. The Weibull survival function is widely used in survival analysis to model the probability of an event not occurring up to a certain time. The standard survival function for the Weibull model is given in Eq. (3.2)

$$S(t;\lambda,k) = e^{(\frac{-t}{\lambda})^k} \tag{3.2}$$

where:

t: The variable of interest (in this case, Post Encroachment Time, PET).

λ: The scale parameter, which determines the spread of PET values.

k: The shape parameter, which determines whether risk increases.

3.9.2 Modified Weibull Survival Function for PET Analysis

In typical survival analysis, survival probability decreases with the time. However, in the context of vehicle-pedestrian interactions, the survival probability exhibits a unique characteristic. As the Post Encroachment Time increases, the probability of a safe interaction also increases, indicating a positive relationship. This behavior is contrary to typical survival analysis, where survival probability decreases over time.

To accommodate this behavior and retain the integrity of the survival curve, the survival function is inverted in this analysis. The Modified Weibull survival function is defined in Eq (3.3):

$$S'(t;\lambda,k) = 1 - S(t;\lambda,k) = 1 - \ e^{(\frac{-t}{\lambda})^k} \tag{3.3}$$

Typically, in survival analysis, time is arranged in ascending order to model cumulative probabilities from the start of the event. However, for PET-specific analysis, descending order is more intuitive because smaller PET values indicate higher risk, and survival probabilities would naturally decrease. The survival probability for the corresponding PET is calculated using Kaplan-Meier survival estimate as shown in Eq. (3.4)

$$S(t_{current}) = S(t_{previous}) * (1 - \frac{m}{n}) \tag{3.4}$$

where,

m is the number of events (conflicts) at the current PET, and

n is the number of interactions at risk just before the current PET.

To model the survival probabilities with corresponding PET values, the Modified Weibull Survival function is employed. Once the Modified Weibull model is fitted to the PET data, the survival function curve is obtained. The second derivative of the fitted curve is analyzed to identify inflection points where the rate of change in survival probability shifts significantly. These inflection points indicate key threshold values that separate different risk regions. Two thresholds are determined to classify risk levels in pedestrian-vehicle interactions. The first threshold corresponds to the peak of the second derivative, indicating the point where survival probability increases most rapidly. PET values below this threshold are categorized as high-risk interactions, as even slight increases in PET significantly reduce the likelihood of a collision. The second threshold is identified where the second derivative stabilizes, signifying diminishing changes in survival probability. PET values beyond this threshold are classified as low-risk interactions, as the temporal gap between pedestrians and vehicles becomes sufficiently safe, reducing the risk of severe conflicts.

3.10 Severity Analysis

In this study, risk assessment primarily relied on PET to identify high-risk vehicle-pedestrian interactions. However, it is essential to distinguish risk (likelihood of conflict) from severity (potential consequences of a collision). While low PET values indicate high-risk interactions, they do not always correlate with high-severity outcomes. For instance, a large vehicle with a higher PET may result in greater severity than a smaller vehicle with a lower PET. Thus, PET alone cannot comprehensively address the severity of conflicts. To enhance the accuracy of severity assessment, additional factors such as vehicle mass (200 kg for 2-Wheelers, 1840 kg for SUV-Cars, 16500 kg for Bus-Trucks) and speed were incorporated into the analysis with mass values based on the most commonly sold vehicle models in each category. Severity was quantified using formulations based on kinetic energy and Post-Encroachment Time to account for these effects as shown in Eq. 3.5.

$$Severity = \frac{K.E_{avg}}{e^{PET}} \tag{3.5}$$

Given that vehicle mass heavily dominates severity, both KE and e^{PET} were normalized by vehicle type to mitigate this disproportionate influence. However, this normalization made it difficult to compare severity across vehicle types, as larger vehicles still inherently cause greater harm. To ensure comparability of severity across different vehicle types, a Crash-Based Modification Factor (CMF) was introduced. This factor was derived from three years of pedestrian crash data, including recorded fatalities, major injuries, and minor injuries(Kc et al., 2024). To quantify the real-world impact of each crash type, the Maximum Abbreviated Injury Scale (MAIS) was applied as shown in Eq 3.6, assigning weights as MAIS 6 for Fatal Injuries, MAIS 3 for Major Injuries and MAIS 1 for Minor Injuries(Stevenson et al., 2001).

$$CMF = \frac{(N_{Fatalities} \times 6) + (N_{MajorInjuries} \times 3) + (N_{MinorInjuries} \times 1)}{Total\ number\ of\ Crashes} \tag{3.6}$$

Finally, the Severity Index were obtained by multiplying the Severity by the Crash-Based Modification Factor as shown in Eq 3.7. Once the Severity Index were derived, they were categorized into Six distinct levels: Very Low, Low, Moderate, High, Very High and Extreme Severity. The classification was determined using percentile-based thresholds

(5th, 25th, 50th, 75th, and 95th percentiles) following a distribution-based approach across the full dataset, consistent with existing literature (Vedagiri & Kadali, 2016;Harkey et al., 1998).

$$Severity\ Index = \frac{K.E_{avg}}{e^{PET}} \times CMF \tag{3.7}$$

CHAPTER 4: RESULTS AND DISCUSSIONS

4.1 Descriptive Analysis

According to Table 4.1, at both sites, the working-age population (20-50 years) is the most dominant, accounting for around 65% of the pedestrians. In Shantinagar, assertive pedestrians, those who confidently claim their right of way while crossing, are significantly more prevalent compared to Dhobighat (61.41% vs. 23.62%). This can be attributed to the high volume of two-wheelers (68.65%) at Shantinagar, which encourages pedestrians to be more assertive to navigate through the dense traffic flow, as these vehicles yield less and rarely give way to pedestrians. In contrast, Dhobighat experiences high-speed traffic with a greater diversity of vehicle types, including a notable proportion of SUV-Cars (34.44%) and Bus-Trucks (14.16%). This composition contributes to a more cautious pedestrian behavior, as people tend to wait for traffic to clear before crossing, leading to a lower percentage of assertive crossings (23.62%) and a higher percentage of reserved behavior (76.37%). Consequently, No Exposure Time (NET) is significantly higher at Dhobighat, as pedestrians prefer to wait for safe gaps rather than expose themselves to oncoming vehicles. Vehicle type plays a crucial role in pedestrian-vehicle interactions at both locations. In Shantinagar, two-wheelers (68.65%) are involved in the majority of conflicts, followed by SUV-Cars (28.28%). In contrast, conflicts involving buses and trucks (3.05%) are rare, as these vehicles primarily use service lanes for passenger pick-up and drop-off. Additionally, heavy vehicles are restricted inside the ring road during daytime, reducing their presence and consequently minimizing conflicts with pedestrians. At Dhobighat, although service lanes are also used by heavy vehicles for pick-up and drop-off, there are no daytime restrictions for these vehicles on the ring road. As a result, there is a significantly higher involvement of buses and trucks (14.16%) in conflicts. This unrestricted access contributes to the increased frequency of heavy vehicle interactions with pedestrians.

The yielding behavior of vehicles is also notably different between the two sites. At Dhobighat, the yielding rate is lower compared to Shantinagar due to the higher speed of vehicles, making drivers less likely to stop for pedestrians. Regarding conflict types, Hybrid Conflicts—where both maneuvering and speed reduction occur simultaneously—are more common at Shantinagar. This is largely due to the narrower roadway, which forces drivers to slow down and navigate carefully. In contrast, at Dhobighat, vehicles either maneuver or reduce speed but rarely do both at once. The presence of high-speed movements and heavy vehicles reduces the occurrence of hybrid conflicts, as the traffic flow is more streamlined and less disrupted by pedestrian crossings.

Table 4.1: Descriptive Analysis of Categorical Variables

Categorical variable	**Composition**	
	Shantinagar	**Dhobighat**
Gender [Male, Female]	[50.4 %, 49.6%]	[52.75 %, 47.25%]
Age [<20, 20-50, >50]	[23.03%, 64.72%, 12.25%]	[20.05%, 65.05%, 14.90%]
Size [Size1, Size2, Size>2]	[58.71%, 25.51%, 15.76%]	[59.87%, 23.62%, 16.51%]
Own Path [Assertive, Reserved]	[61.41%, 38.59%]	[23.62%, 76.37%]
Crossing Pattern [Jaywalker, Normal]	[18.80%, 81.95%]	[19.42%, 80.58%]
Stage of Crossing [Single, Multiple]	[72.61%, 27.39%]	[72.81%, 27.18%]
Curb Time [NWT, SWT, LWT]	[65.97 %, 17.84 %, 16.182%]	[71.19 %, 20.71 %, 8.10%]
Exposure Time [NET, SET, LET]	[48.34 %, 32.57% ,19.08%]	[62.78 %, 20.06% ,17.15%]
Conflict Location [Near Lane, Fare Lane]	[44.64%, 55.36%]	[24.31%,75.69%]
Vehicle Type [2W, Suv-Car, Bus-Truck]	[68.65%, 28.28%, 3.05 %]	[51.40%, 34.44%, 14.16 %]
Yield [Yes, No]	[17.73%, 82.27%]	[12.50%, 87.50%]
Conflict Type [Maneuver Conflict, Speed based Conflict, Hybrid Conflict]	[39.44%, 37.00%, 23.54%]	[40.27%, 47.22%, 12.50%]

Table 4.2 shows the descriptive analysis of continuous variables at Shantinagar and Dhobighat. The PET values are consistently higher at Dhobighat. This cautious behavior is likely due to the higher traffic speeds and diverse vehicle types at Dhobighat. In contrast, the lower PET values at Shantinagar suggest more assertive crossing behavior, influenced by the dominance of two-wheelers, which pedestrians perceive as less threatening. Traffic speeds are higher at Dhobighat across all percentiles due to the high-speed environment. In contrast, Shantinagar experiences lower speeds, likely due to the dense flow of vehicles and the urban setting, which limits high-speed movement. Among different vehicle types, two-wheelers exhibit the highest speeds at both locations, taking advantage of their maneuverability, while larger vehicles such as buses and trucks maintain more moderate speeds due to their size and stopping constraints. This speed difference influences pedestrian crossing behavior, as seen in the PET values. Shantinagar consistently shows a higher perceived traffic volume compared to Dhobighat, primarily due to the dense flow of two-wheelers and the urban nature of the area. Pedestrians at Shantinagar cross the road at slightly higher speeds compared to those at Dhobighat. This is likely due to the higher traffic volume at Shantinagar, prompting pedestrians to take advantage of gaps in traffic and cross quickly. In contrast, at Dhobighat, the presence of large gaps due to less volume leads to more cautious and slightly slower crossing speeds. Across both sites, the <20 age group exhibits the highest crossing speeds, while those over 50 cross the slowest, reflecting differences in agility and risk perception.

Table 4.2: Descriptive Analysis of Continuous Variables

Continuous Variable		Shantinagar			Dhobighat		
		Q1	Median	Q3	Q1	Median	Q3
PET (s)		1.00	1.27	1.57	1.46	1.76	2.11
Traffic Speed (m/s)	2W	7.38	9.21	11.24	8.33	10.11	12.01
	SUV-Car	5.80	7.35	9.79	7.83	9.45	11.27
	Bus-Truck	6.05	7.71	8.08	8.07	9.05	11.51
Perceived Vol (veh/min)		42	59	80	18	30	48
Crossing Speed (m/s)	<20	1.10	1.20	1.34	1.12	1.21	1.33
	20-50	1.08	1.19	1.31	1.04	1.16	1.30
	>50	0.95	1.10	1.20	0.97	1.06	1.14

4.2 Correlation Analysis

A Pearson correlation analysis was conducted to examine the relationships among the continuous variables in the dataset, while Cramér's V test was used to assess associations between categorical variables. The Pearson correlation results as shown in Fig 4.1 indicate that all correlation values among the independent variables are below 0.5, suggesting no significant multicollinearity. This confirms that each variable independently contributes to the analysis without redundancy.

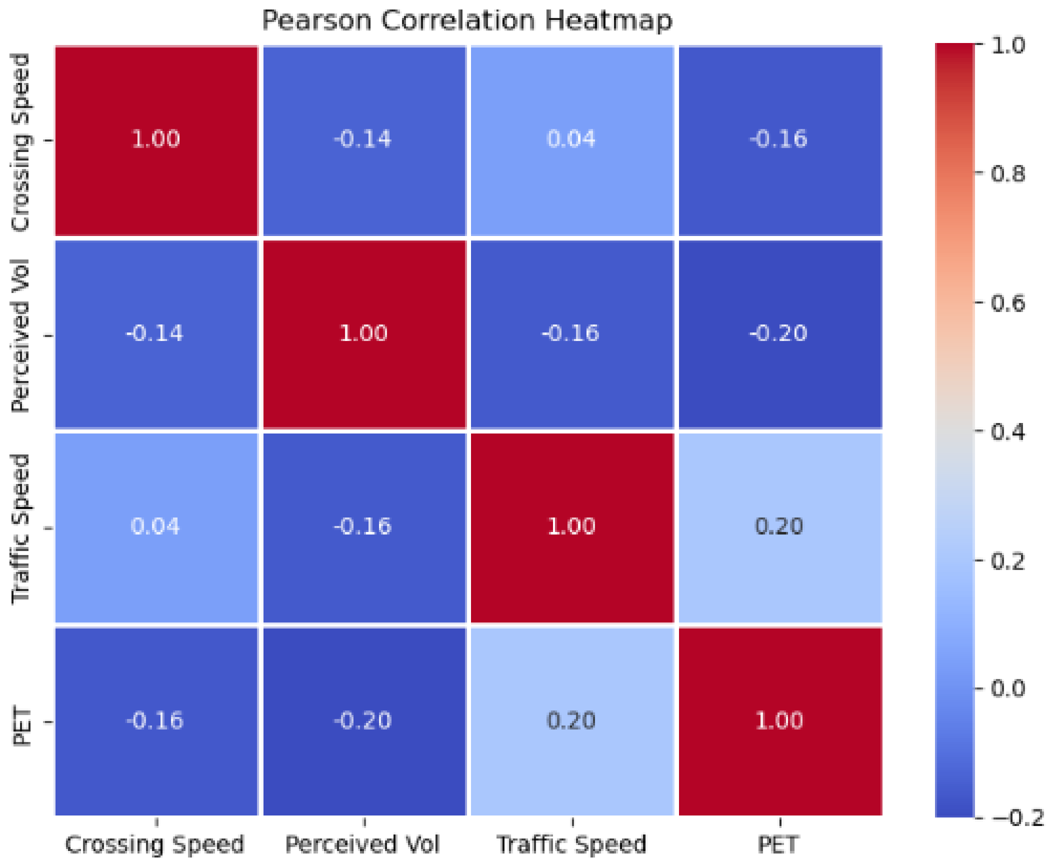


Figure 4.1 : Pearson Correlation Heat Map for Continuous Variables

Similarly, in the Cramér's V test, as shown in Fig 4.2, a notable correlation (0.66) was observed between Exposure Time and Crossing Stage, indicating a moderate association. Additionally, a moderate correlation (0.37) was found between Yield and Conflict Type, suggesting a relationship between these variables. A lower but noticeable correlation (0.30) was observed between Vehicle Type and Conflict Type. However, multicollinearity is typically assessed using Variance Inflation Factor (VIF) in regression modeling as shown in Table 4.3

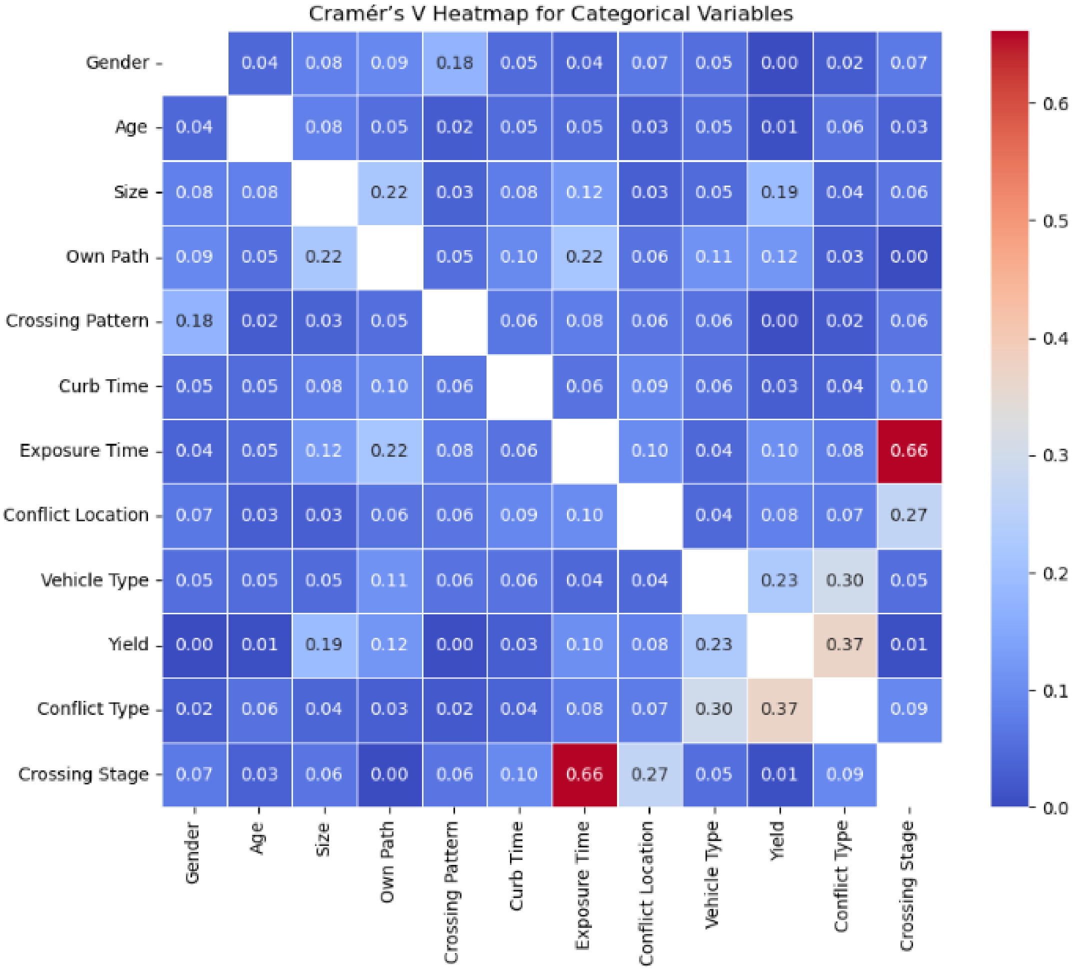


Figure 4.2: Cramer's V Heatmap for Categorical Variables

Therefore, all variables were retained in the initial regression analysis to evaluate their individual contributions to PET.

Table 4.3: Features with Corresponding VIF

Feature	**VIF**
ordinal__Age	1.07
ordinal__Size	1.21
ordinal__Curb Time	1.16
ordinal__Exposure Time	1.81
nominal__Gender_Male	1.10

Feature	VIF
nominal__Own Path_Reserved	1.21
nominal__Crossing Pattern_Normal	1.05
nominal__Vehicle Type_Bus-Truck	1.19
nominal__Vehicle Type_SUV-Car	1.24
nominal__Yield_Yes	1.41
nominal__Crossing Stage_Single	1.87
nominal__Conflict Location_Near Lane	1.13
nominal__Conflict Type_Manuever Conflict	2.01
nominal__Conflict Type_Speed Based Conflict	2.09
numeric__Traffic Speed	1.45
numeric__Crossing Speed	1.22
numeric__Perceived Vol	1.39

4.3 Multiple Linear Regression Model for Post Encroachment Time

An Ordinary Least Squares (OLS) regression analysis was conducted to model Post Encroachment Time (PET) using a combination of categorical and numerical predictors. In Table 4.4, the initial model explained 29% of the variance in PET ($R^2 = 0.29$), with an adjusted R^2 of 0.277, indicating a moderate level of explanatory power. The model's overall significance was confirmed by an F-statistic of 22.21 ($p = 1.84E-57$), demonstrating that the predictors collectively contribute to explaining PET variations. The model's log-likelihood was -488.79, with AIC and BIC values of 1014 and 1101, respectively, providing insights into model fit and complexity.

The p-values of individual predictors, as reported in Table 4.5, were used to determine their statistical significance. Variables with p-values less than 0.05 were considered significant contributors to PET. Key significant predictors included Curb Time ($p = 0.042$), Own Path - Reserved ($p = 0.002$), Vehicle Type - Bus/Truck ($p = 0.000$), Vehicle Type - SUV/Car ($p = 0.000$), Yield - Yes ($p = 0.001$), Crossing Stage - Single ($p = 0.041$), Conflict Type - Maneuver Conflict ($p = 0.000$), Conflict Type - Speed-Based Conflict ($p = 0.000$), Traffic Speed ($p = 0.000$), Crossing Speed ($p = 0.000$), and Perceived Volume ($p = 0.000$).

Conversely, some variables, such as Age (p = 0.197), Size (p = 0.571), Exposure Time (p = 0.191), and Gender (p = 0.453), were found to be statistically insignificant, suggesting they do not have a strong influence on PET.

Table 4.4: OLS Regression Results of Model Considering All Variables.

OLS Regression Results						
R-squared	Adj. R-squared	F-Statistic	Prob(F- statistic)	Log-Likelihood	AIC:	BIC:
0.29	0.277	22.21	1.84E-57	-488.79	1014	1101

Table 4.5: Features with Corresponding P-values

Feature	**p-values**
const	0.000
ordinal__Age	0.197
ordinal__Size	0.571
ordinal__Curb Time	0.042
ordinal__Exposure Time	0.191
nominal__Gender_Male	0.453
nominal__Own Path_Reserved	0.002
nominal__Crossing Pattern_Normal	0.137
nominal__Vehicle Type_Bus-Truck	0.000
nominal__Vehicle Type_SUV-Car	0.000
nominal__Yield_Yes	0.001
nominal__Crossing Stage_Single	0.041
nominal__Conflict Location_Near Lane	0.095
nominal__Conflict Type_Manuever Conflict	0.000
nominal__Conflict Type_Speed Based Conflict	0.000
numeric__Traffic Speed	0.000
numeric__Crossing Speed	0.000
numeric__Perceived Vol	0.000

To improve the model’s interpretability and predictive performance, a refined model was developed, retaining only the significant predictors. The results of the refined model are presented in Table 4.6 and 4.7

Table 4.6: OLS Regression Results of Model Considering Significant Variables.

OLS Regression Results						
R-squared	Adj. R-squared	F-Statistic	Prob(F- statistic)	Log-Likelihood	AIC:	BIC:
0.283	0.274	33.32	5.62E-60	-493.67	1011	1070

Table 4.7: Features with All Parameters.

Feature	Coeff β	Standard Errors	t-values	p-values
const	1.12	0.04	27.06	0.00
ordinal__Curb Time	0.04	0.02	1.94	0.05
nominal__Own Path_Reserved	0.08	0.03	2.97	0.00
nominal__Vehicle Type_Bus-Truck	0.42	0.06	7.09	0.00
nominal__Vehicle Type_SUV-Car	0.20	0.03	6.28	0.00
nominal__Yield_Yes	0.14	0.04	3.32	0.00
nominal__Crossing Stage_Single	-0.10	0.03	-3.09	0.00
nominal__Conflict Type_Manuever Conflict	0.40	0.04	10.31	0.00
nominal__Conflict Type_Speed Based Conflict	0.24	0.04	6.15	0.00
numeric__Traffic Speed	0.08	0.02	4.79	0.00
numeric__Crossing Speed	-0.08	0.01	-5.77	0.00
numeric__Perceived Vol	-0.08	0.02	-5.32	0.00

Table 4.7 illustrates that the Vehicle type significantly impacted PET. The interactions with Bus-Trucks ($\beta = 0.42$, $p < 0.001$) and SUV-Cars ($\beta = 0.20$, $p < 0.001$) were associated with longer PET. This could be due to the larger physical size of these vehicles, making pedestrians more cautious and prompting them to wait longer before crossing or adjust their movements more carefully to avoid conflicts.

Maneuver conflicts ($\beta = 0.40$, $p < 0.001$) and speed-based conflicts ($\beta = 0.24$, $p < 0.001$) were both associated with increased PET. In maneuver conflicts, either the vehicle or the pedestrian may maneuver laterally over a larger distance. This extended lateral movement effectively delays the point at which the two parties meet, thereby increasing the time until the conflict is resolved. Similarly, in speed-based conflicts, one or both parties might adjust their speed, either by slowing down to delay reaching the conflict point, or by accelerating to move away more quickly which also results in a longer duration before the conflict region is reached.

Traffic speed ($\beta = 0.08$, $p < 0.001$) was positively correlated with PET, while pedestrian crossing speed ($\beta = -0.08$, $p < 0.001$) had a negative association. Higher traffic speeds may cause pedestrians to be more cautious, increasing PET. Certain pedestrian behaviors also influenced PET. Longer curb waiting time ($\beta = 0.04$, $p = 0.05$) was positively associated with PET, suggesting that cautious pedestrians who wait longer before crossing may experience prolonged interactions. Similarly, reserved pedestrians ($\beta = 0.08$, $p = 0.002$) tend to wait at the curb and refrain from using hand gesture to expedite their crossing. This behavior suggests that they only initiate crossing when conditions appear safest, which, while reducing risk, increases the overall interaction time with vehicles (PET).

Yielding behavior by vehicles ($\beta = 0.14$, $p < 0.001$) was also associated with increased PET. When drivers yield to pedestrians, it introduces a delay as pedestrians confirm the vehicle's intent before proceeding, thereby prolonging PET. Higher perceived traffic volume ($\beta = -0.08$, $p < 0.001$) was significantly linked to a reduction in Post Encroachment Time (PET). This effect may be attributed to the dominance of two-wheelers in traffic, which often creates narrower and more frequent gaps between vehicles. Pedestrians may perceive two-wheelers as less threatening compared to larger vehicles, leading them to take riskier crossings through these smaller gaps, especially under the urgency created by heavy traffic conditions.

The refined model, which retains only significant variables, demonstrates a better balance between model fit and simplicity. It has lower AIC (1011) and BIC (1070) compared to the full model (1014 and 1101), with a difference of 3 in AIC and 31 in BIC—strongly favoring the reduced model, as a difference greater than 2 is considered meaningful. Additionally, the reduced model has a higher F-statistic (33.32 vs. 22.21), indicating that it explains the variance in PET more efficiently relative to the number of predictors used. While the full model has a slightly better log-likelihood (-488.79 vs. -493.67), the difference is minor, and since AIC and BIC penalize unnecessary complexity, the reduced model remains preferable. The adjusted $R^2 = 0.27$ remains nearly the same across both models, indicating that the simpler model retains similar explanatory power while avoiding unnecessary complexity. In behavioral analysis, low R^2 values are common due to the inherent variability and complexity of human behavior, and such values have been reported in previous studies as well (Kadali & Vedagiri, 2016 ;Piyalungka et al., 2025). In empirical research, particularly within the social sciences, a low R-square value does not necessarily indicate a poor model. The primary goal of such research is often not to predict outcomes with high precision, but to assess the significance of specific predictors or explanatory variables in influencing the dependent variable. Therefore, an R-square value as low as 0.1 (10%) can still be acceptable, provided that the predictors or explanatory variables are statistically significant (Ozili, 2022).

4.4 Multiple Linear Regression Model Validation

Model validation ensures that an MLR model is statistically sound and reliable for predictions. It involves checking key assumptions and performance metrics. As we can see in the Figure 4.3, the residual histogram approximates a normal distribution, and the Q-Q plot confirms normality.

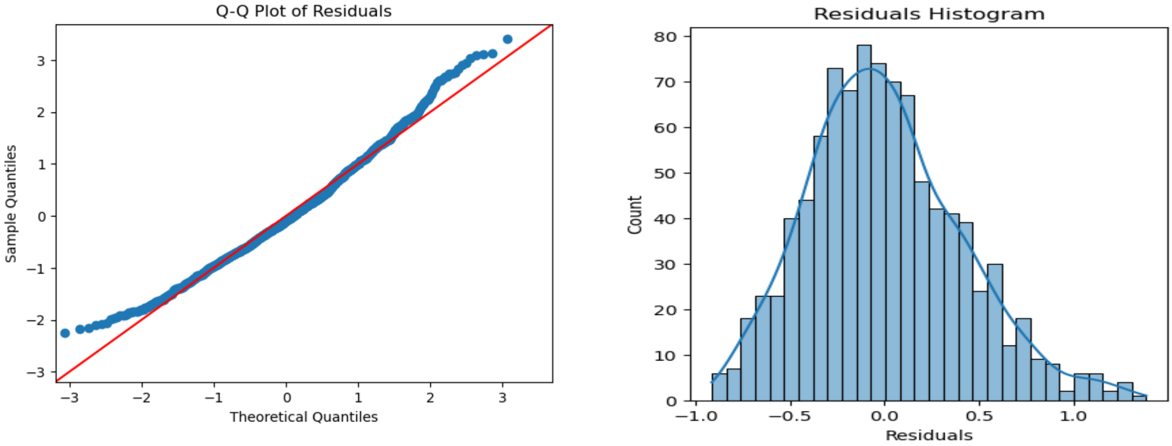


Figure 4.3 : Distribution of Residuals and Q-Q Plots

The residuals vs fitted values are plotted in Fig 4.4 which shows a random scatter, indicating no heteroscedasticity. Multicollinearity was evaluated using the Variance Inflation Factor (VIF), with all values below 2.1 which is shown in Table 4.3, confirming no severe multicollinearity. The Durbin-Watson test result of 1.795 indicates minimal autocorrelation, ensuring the independence of residuals. These validations confirm that the model is statistically sound and reliable for predictions.

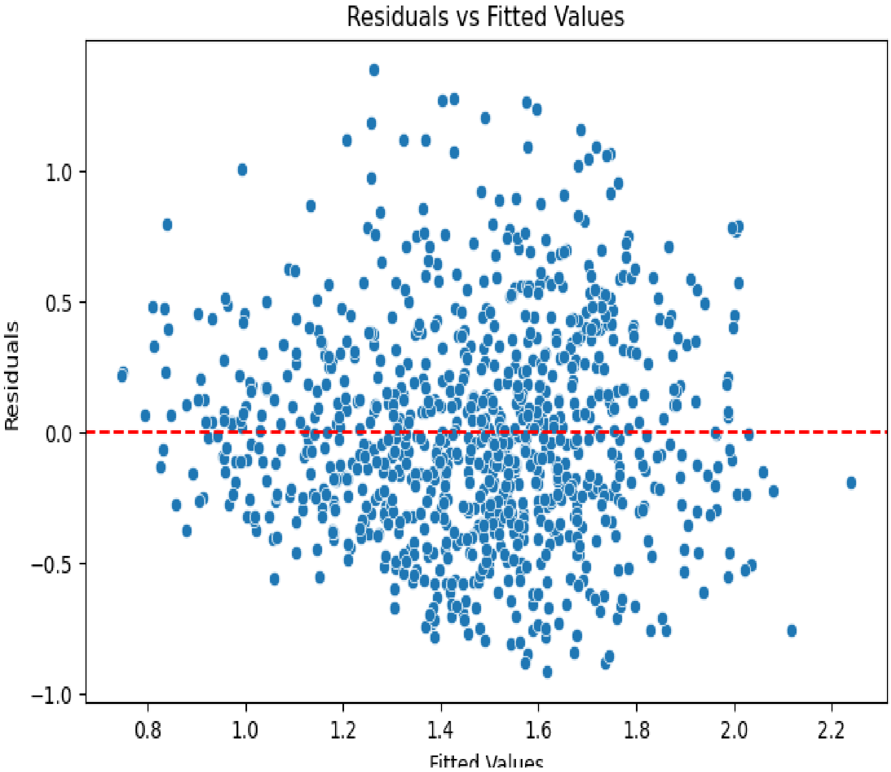


Figure 4.4: Plot of Residuals vs Fitted Values

4.5 Modified Weibull Survival Model

The Modified Weibull Survival Function was successfully fitted to the survival probability data. In Figure 4.5, the blue dots represent the observed data points of PET values and their corresponding survival probabilities, while the red curve shows the fitted Modified Weibull Survival Function (also called Fitted Weibull CDF) which closely follows the distribution of the data. The smooth progression of the CDF illustrates the cumulative likelihood of survival as PET increases. In traditional Weibull analysis, a shape parameter $k > 1$ indicates that the hazard rate (risk of an event) increases over time. However, in Modified Weibull function, $k > 1$ suggests that as PET increases, the risk of conflict decreases over time. From the model, the estimated Weibull parameters are shape parameter ($k = 3.24$) and scale parameter ($\lambda = 1.35$) which align with expected risk dynamics. Also, the visual alignment as show in Figure 4.5 between the observed data points and the Fitted Weibull CDF confirms the model validity.

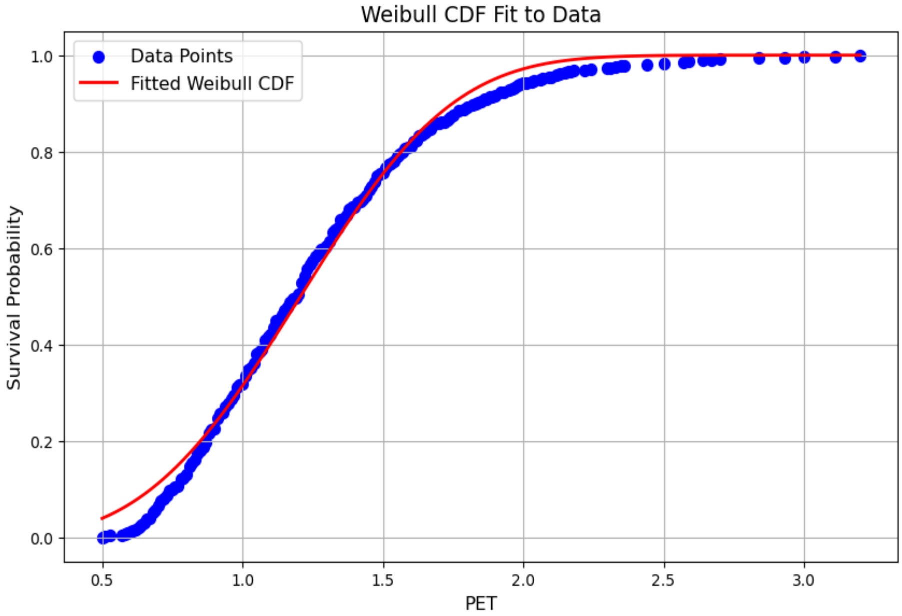


Figure 4.5: Modified Weibull Survival Function Fit to Data of Shantinagar

In order to analyze risk levels, the second derivative of the fitted Weibull CDF was computed, providing the rate of change of the slope. This indicates that even slight increases in PET values result in a significant rise in the survival probability, helping to pinpoint key thresholds where risk level shift. The Figure 4.6 visualizes the second derivative (rate of change of slope) of the fitted Weibull CDF across PET values.

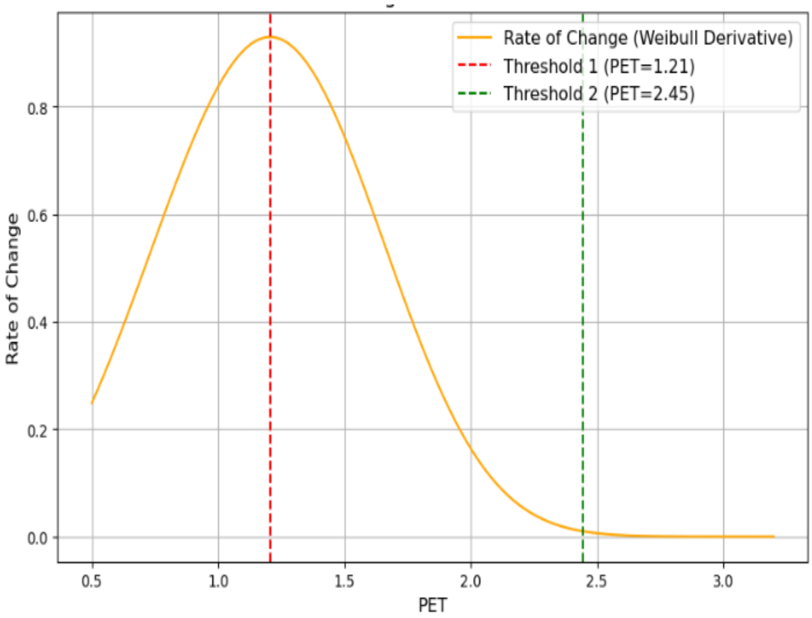


Figure 4.6 : Rate of Change of Slope of Fitted Weibull CDF of Shantinagar

The curve peaks at PET = 1.21 (Threshold 1), indicating the steepest increase in survival probability. At this point, even small increases in PET lead to the most significant reduction in collision risk. Beyond PET = 2.45 (Threshold 2), the rate of change flattens, signaling that survival probability stabilizes, and further increases in PET result in diminishing improvements in safety.

The Figure 4.7 illustrates the classification of risk regions for vehicle-pedestrian interactions based on Post Encroachment Time (PET) using a fitted Weibull Cumulative Distribution Function (CDF). The graph shows the relationship between PET and survival probability, highlighting three distinct risk regions. The High-Risk Region (yellow area) is located to the left of the first threshold (red dashed line) at PET = 1.21 seconds. It corresponds to low PET values associated with high-risk interactions, where the likelihood of a severe incident is greatest due to the short temporal distance between the pedestrian and vehicle. Medium-Risk Region (cyan area) lies between the two thresholds, spanning from PET = 1.21 to PET = 2.45 seconds. It represents moderate risk, where the temporal gap is sufficient to reduce the risk of severe consequences but still poses a potential hazard. Low-Risk Region (purple area) is situated to the right of the second threshold (green dashed line) at PET = 2.45 seconds, indicating low-risk interactions as the PET values are high enough to ensure a safer separation between pedestrian and vehicle.

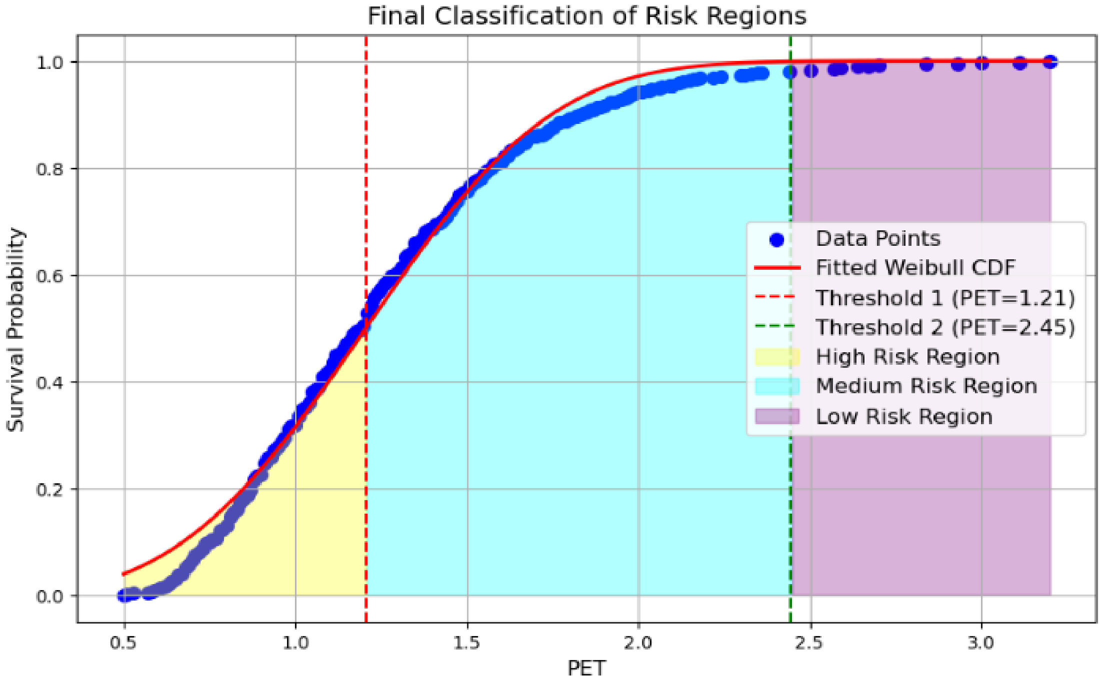


Figure 4.7: Classification of Risk Regions based on Threshold PET Values for

A similar analysis was performed for Dhobighat using the same Modified Weibull Survival Function. The estimated shape and scale parameters were (k = 3.97), (λ = 1.84), aligning with the risk dynamics observed in Shantinagar but reflect subtle differences in interaction patterns. In comparison to Shantinagar (k =3.24, λ =1.35), Dhobighat exhibits a higher shape parameter (k), suggesting an even steeper decline in collision risk as PET increases. This implies that in Dhobighat, marginal improvements in PET lead to a more pronounced reduction in hazard rates compared to Shantinagar. The scale parameter λ=1.84, reflects that survival probability stabilizes at larger PET values.

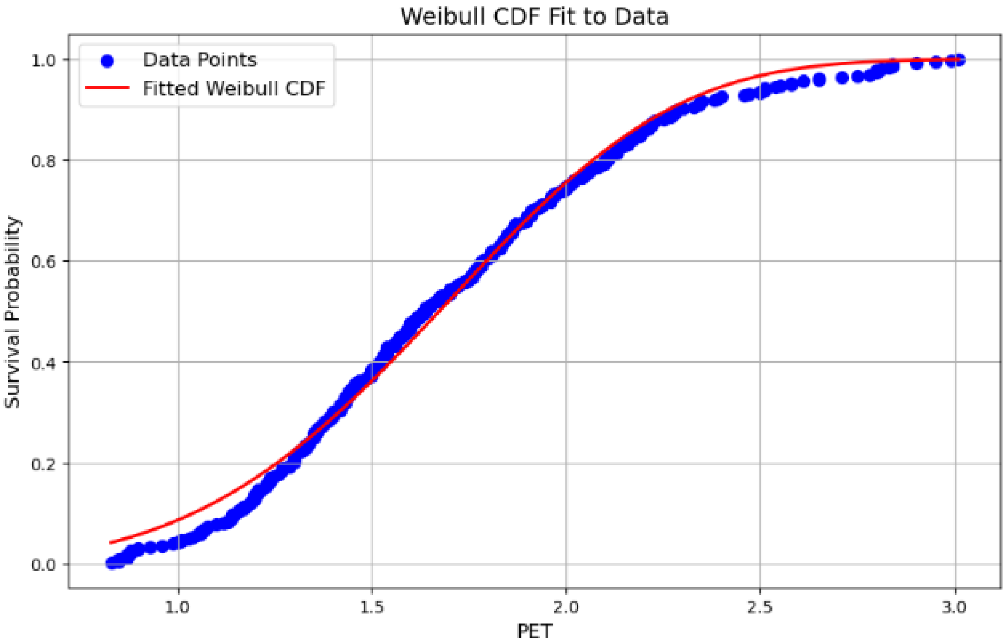


Figure 4.8: Modified Weibull Survival Function Fit to Data of Dhobighat

The key PET thresholds were identified at PET = 1.71 (Threshold 1) and PET = 2.98 (Threshold 2), marking the transition between risk regions. These threshold values are higher than those observed at Shantinagar (1.21s and 2.45s, respectively), indicating that pedestrians at Dhobighat experience longer PET values before reaching similar risk transition points. This can be attributed to the distinct traffic characteristics at Dhobighat, where higher vehicle speeds and the presence of heavy vehicles influence pedestrian behavior. Unlike Shantinagar, where pedestrians exhibit more assertive crossings due to the dominance of two-wheelers, Dhobighat's traffic environment encourages more cautious pedestrian behavior. Despite the longer PET thresholds, the risk dynamics remain critical, as even at PET = 1.71s, pedestrians may still face high-risk interactions due to the high-speed environment and the presence of larger vehicles with restricted maneuverability, making evasive action more challenging.

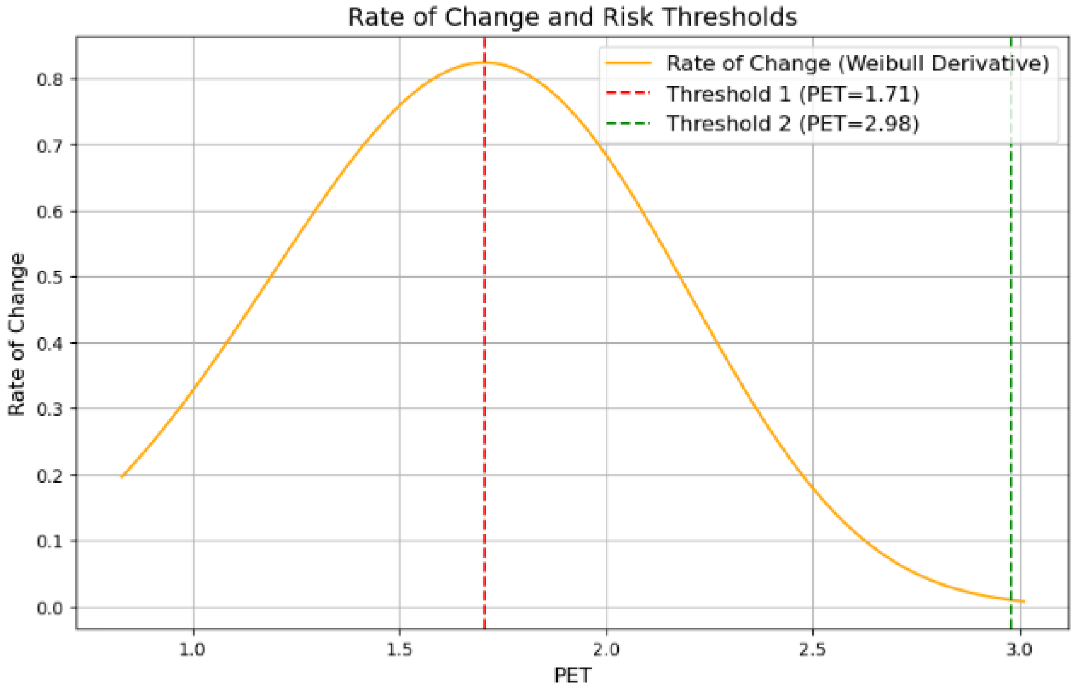


Figure 4.9: Rate of Change of Slope of Fitted Weibull CDF of Dhobighat

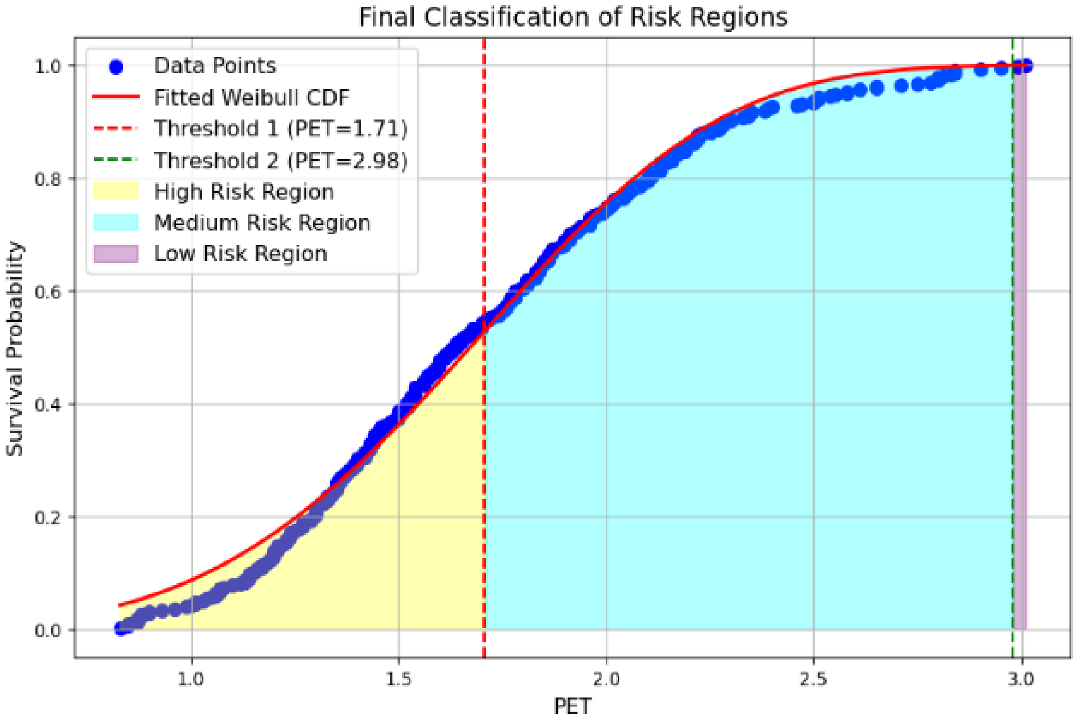


Figure 4.10: Classification of Risk Regions for Dhobighat

To establish a unified threshold value for mixed traffic conditions in Nepal, the datasets from two distinct locations were combined, and the Modified Weibull Survival Function was fitted to the aggregated Post Encroachment Time data as shown in Fig. 4.11.

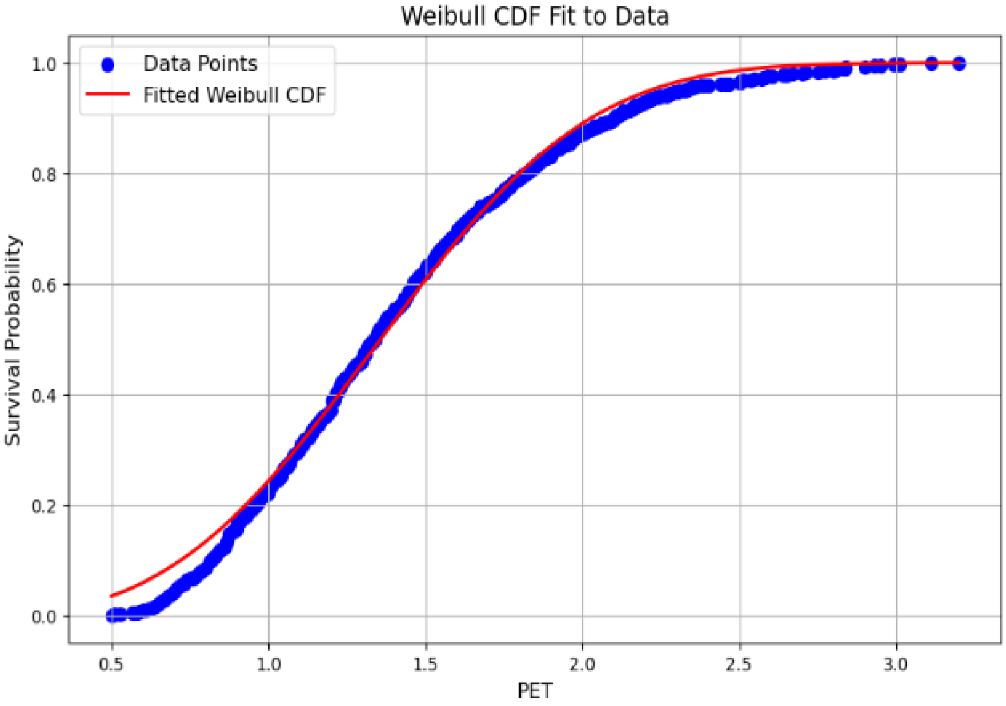


Figure 4.11: Modified Weibull CDF Fit to Combined Datasets

This analysis yielded a shape parameter k=2.99 and a scale parameter λ=1.54, providing insight into the risk dynamics of vehicle-pedestrian interactions across the combined dataset. The shape parameter k=2.99, being greater than 1, indicates that the risk of conflict decreases as PET increases, with a moderately steep decline in hazard rate reflective of the mixed traffic environment. The scale parameter λ=1.54 represents the characteristic PET value at which approximately 63.2% of the survival probability is achieved, suggesting a balanced spread of PET values within the dataset. The fitted Weibull Cumulative Distribution Function (CDF) effectively captures the cumulative likelihood of survival, aligning with the observed data and validating the model's applicability to Nepal’s heterogeneous traffic context. To further analyze risk levels, the second derivative of the fitted CDF was computed, revealing two critical thresholds. From Fig 4.12, Threshold 1, at PET = 1.34 seconds, marks the point where the survival probability increases most rapidly, indicating that small increases in PET beyond this value lead to significant reductions in collision risk.

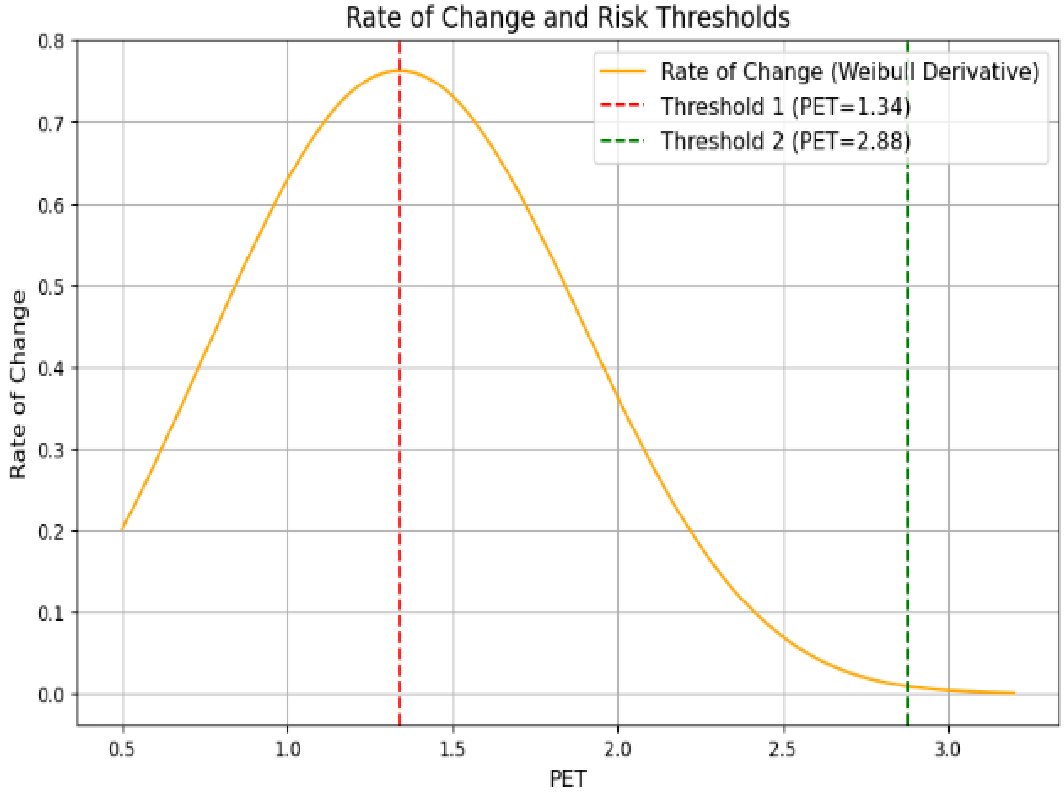


Figure 4.12: Rate of Change of Slope of Fitted Weibull CDF of Combined Datasets

This threshold delineates the boundary of the High-Risk Region, where PET values below 1.34 seconds correspond to heightened danger due to insufficient temporal separation between vehicles and pedestrians. Threshold 2, at PET = 2.88 seconds, represents the point where the rate of change in survival probability begins to flatten, signaling that further increases in PET yield diminishing improvements in safety. Beyond this threshold, interactions enter the Low-Risk Region, where the temporal gap is sufficient to ensure minimal risk. In Fig 4.13, these thresholds—1.34 seconds and 2.88 seconds—define three risk regions: a High-Risk Region below 1.34 seconds, a Medium-Risk Region between 1.34 and 2.88 seconds, and a Low-Risk Region above 2.88 seconds. In general, PET values below 1 second are considered critical, whereas PET values greater than 2 seconds are typically associated with normal, lower-risk encounters (VAN DER HORST, 1990). However, in our case, the PET thresholds are comparatively higher, likely due to the influence of mixed traffic conditions, varied vehicle types, and differing pedestrian behaviors. The presence of a heterogeneous traffic environment, with a mix of slow-moving and fast-moving vehicles, restricted maneuverability, and varying yielding behaviors, may contribute to longer PET values before transitioning between risk levels.

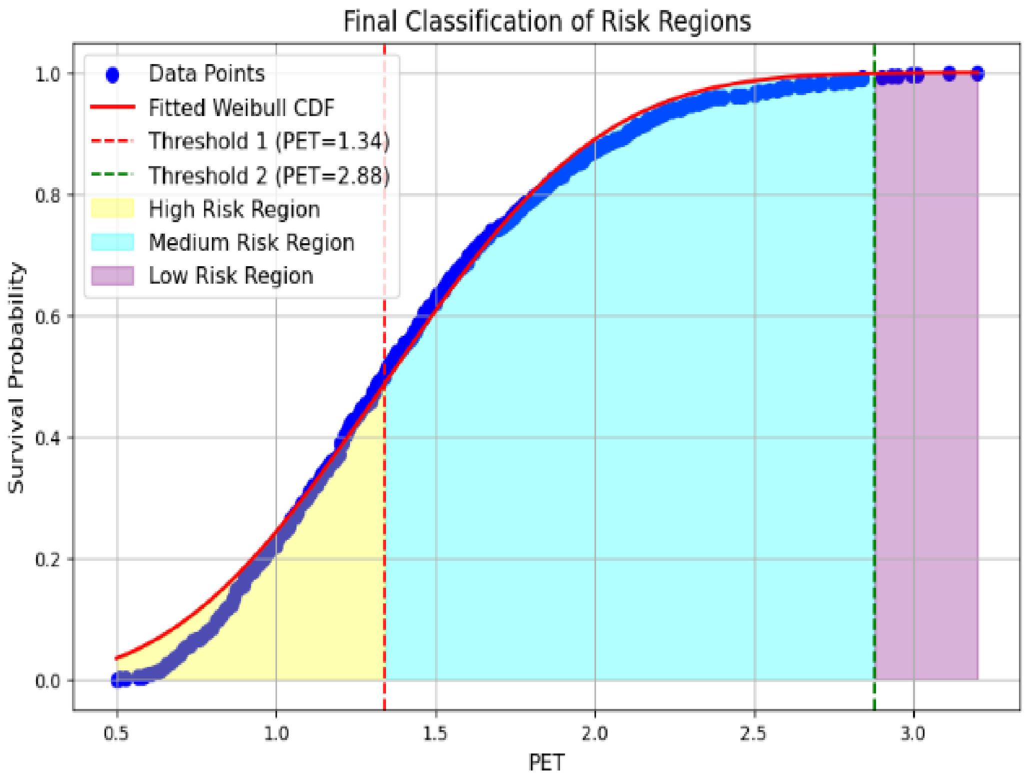


Figure 4.13: Classification of Risk Regions for Combined Datasets

4.6 Severity Analysis

The severity analysis was conducted for both study locations, incorporating kinetic energy (KE), post-encroachment time (PET), and a Crash-Based Modification Factor (CMF) to account for real-world injury severity variations. Since vehicle mass strongly influences KE, direct comparisons across vehicle types were challenging. To address this, Min-Max normalization was applied separately to KE and e^{PET} for each vehicle category, ensuring that normalized values ranged between 0 and 1 while preserving relative severity differences. Once normalized values were obtained, the Severity Index (SI) was computed by incorporating the CMF, which was derived from pedestrian crash data using a weighted scale for different injury severities. Table 4.8 shows the resulting CMF values were 1.25 for Two-Wheelers, 1.49 for SUV- Cars and 2.11 for Buses/Trucks. To understand the distribution of the Severity Index across all vehicle types and within each vehicle category, Kernel Density Estimation (KDE) plots were generated. Figure 4.14 presents the KDE plot of the Severity Index for all vehicle types combined, revealing a right-skewed distribution with a peak around 1.5 and a long tail extending to approximately 20.

Table 4.8: Calculation of Crash Based Modification Factor

Severity Level	MAIS Scale	2 W	SUV-Car	Bus-Truck
Death	6	35	18	26
Major Injury	3	66	30	22
Minor Injury	1	1204	274	110
Number of Crashes		1292	317	157
Crash Based Modification Factor		1.25	1.49	2.11
Fatalities %		3 %	6 %	17 %

This indicates that the majority of pedestrian interactions resulted in relatively moderate to high severity, but a small proportion of interactions exhibited extreme severity, likely driven by crashes involving heavier vehicles or higher speeds. The distribution is right-skewed, indicating that most interactions have relatively low severity values, with fewer interactions exhibiting high severity. The density is highest near moderate severity values, suggesting that a majority of pedestrian-vehicle interactions result in lower severity outcomes, but the presence of a long tail signifies the existence of high-severity incidents.

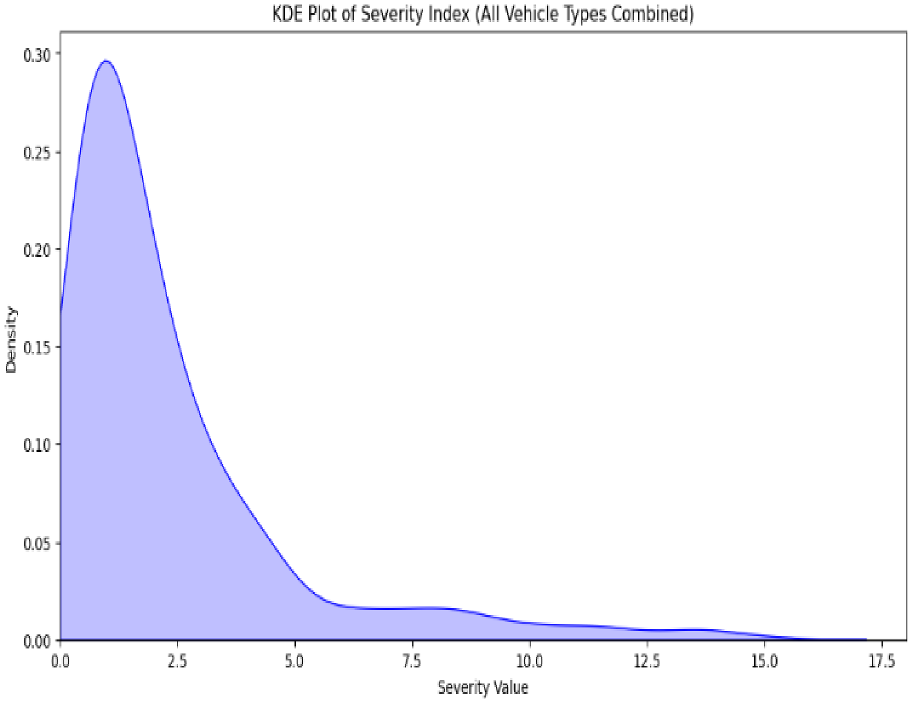


Figure 4.14: KDE Plot of Severity Index for Pedestrian Interaction with all Vehicle Types

Figure 4.15 illustrates the KDE plot of the Severity Index for different vehicle types (2-Wheelers, SUV-Cars, Bus-Trucks) with independent normalization and incorporating CMF, allowing for a direct comparison of the severity distributions. The 2-Wheelers (labeled as "Bike") exhibited a narrow peak at approximately 1, indicating a consistent and predominantly low severity profile, which aligns with their lower mass (180 kg) and the resulting lower kinetic energy in collisions. SUV-Cars displayed a moderate peak at around 1.5, with a wider spread, suggesting a mix medium and high severity interactions, consistent with their intermediate mass (1840 kg) and the potential for higher speeds. Bus-Trucks, with a peak at approximately 2 and the widest distribution, demonstrated the highest average severity and the greatest variability, reflecting the significant impact of their large mass (16,500 kg) and the potential for more severe outcomes in pedestrian crashes.

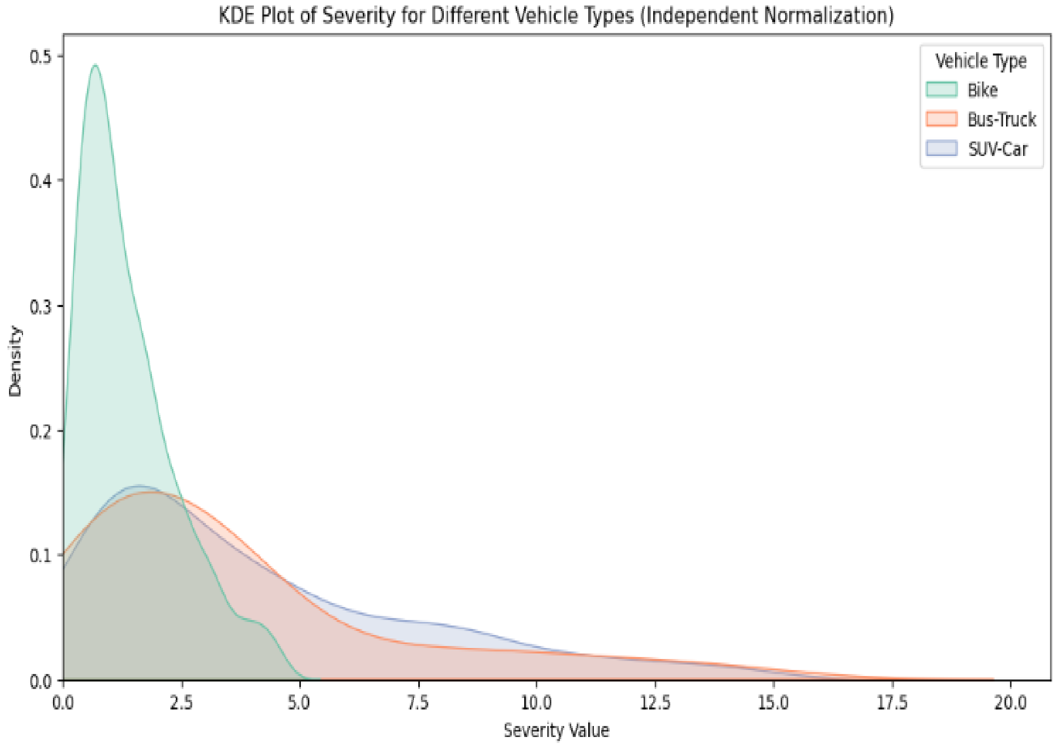


Figure 4.15: KDE Plot of Severity for Pedestrian Interaction with Different Vehicle Types

To further analyze the severity of pedestrian interactions, the Severity Index values were categorized into six levels—Very Low, Low, Moderate, High, Very High, and Extreme—based on the percentiles derived from the dataset. The thresholds were set at the 5th (0.25), 25th (0.79), 50th (1.66), 75th (3.47), and 95th (11.16) percentiles. Accordingly, interactions were classified as follows: Very Low Severity (Severity Index ≤ 0.25), Low Severity (0.25

< Severity Index ≤ 0.79), Moderate Severity (0.79 < Severity Index ≤ 1.66), High Severity (1.66 < Severity Index ≤ 3.47), Very High Severity (3.47 < Severity Index ≤ 11.16), and Extreme Severity (Severity Index > 11.16). The percentage distribution of these severity levels was then computed for each vehicle type to evaluate their risk profiles.

As illustrated in Table 4.9, for 2-Wheelers (Bike), the majority of interactions (26.50% and 29.50%) fell within the Low and Moderate Severity categories, respectively, indicating that most conflicts involved relatively lower impact levels due to their lower mass and kinetic energy. However, 25.87% of interactions reached the High Severity category, and 11.51% were categorized as Very High, suggesting a non-negligible risk of serious pedestrian interactions.

Table 4.9: Percentage Distribution of Severity Levels for Each Vehicle Type

Vehicle Type	Severity Level					
	Very Low	Low	Moderate	High	Very High	Extreme
2 W	6.30	26.50	29.50	25.87	11.51	0.31
SUV-Car	2.58	9.70	16.50	23.62	35.60	11.97
Bus-Truck	4.22	8.45	19.71	22.53	28.16	16.90

The proportion of Extreme Severity cases was minimal (0.32%), reaffirming the limited likelihood of highly severe outcomes with two-wheelers. In contrast, SUV-Cars demonstrated a greater proportion of high-risk interactions, with 23.62% of cases classified as High Severity, 35.60% as Very High, and 11.97% as Extreme. The increased severity risk aligns with their greater mass (~1840 kg) and potential for higher speeds, which contribute to more severe pedestrian conflicts. Bus-Trucks exhibited the highest severity risk among all vehicle types, with 16.90% of cases falling under the Extreme Severity category and 28.17% classified as Very High Severity. The substantial mass (~16,500 kg) and larger impact forces of buses and trucks significantly increase the severity of pedestrian interactions. Notably, only 4.23% of interactions were categorized as Very Low Severity, highlighting the limited number of minor interactions involving heavy vehicles. The overall distribution confirms that as vehicle mass increases, the likelihood of severe pedestrian interactions rises sharply, reinforcing the critical role of vehicle type in determining the severity of pedestrian-vehicle conflict.

To draw a meaningful link between the computed Severity Index and real-world outcomes, fatality data was used for comparison, as fatalities represent the most severe outcome in crash scenarios. In the context of Nepal, this choice is also supported by the fact that fatal crashes are more reliably recorded compared to non-fatal injuries. Due to inconsistencies in how minor and major injuries are reported—often because of unclear hospital classification—relying on injury data could introduce errors. In contrast, fatalities are better documented through legal and institutional processes, making them a more dependable reference point for comparing with the highest severity level i.e. Extreme Severity, which represents the most dangerous interactions in this study. Conceptually, these extreme interactions may correspond to real-world scenarios that are most likely to result in fatalities. Therefore, comparing the proportion of Extreme Severity interactions to the fatality rate across different vehicle types offers a meaningful way to assess the consistency between observed behavior-based severity and actual crash outcomes.

The results of this comparison are summarized based on Table 4.8 and Table 4.9: for two-wheelers, 0.31% of interactions were categorized as Extreme Severity, while 3% of crashes involving two-wheelers resulted in fatalities. For SUV-Cars, 11.97% of interactions fell under Extreme Severity, compared to a 6% fatal crash rate. In the case of buses and trucks, 16.90% of interactions were classified as Extreme Severity, aligning closely with a fatal crash rate of 17%. Although the computed severity levels and observed fatality rates differ in their absolute percentages, the overall trend remains consistent—heavier vehicle types are associated with higher severity in both modeled interactions and real-world crashes. This consistency supports the relevance of the Severity Index in capturing real-world risk dynamics, particularly in distinguishing the danger posed by different vehicle categories. However, it is important to clarify that this comparison is not a formal validation. Since the same crash data was used in part to derive the Crash-Based Modification Factor (CMF), and also serves as the source for fatality rates, there is a risk of methodological overlap. Furthermore, the Severity Index is based on pre-crash behavioral interactions, whereas crash outcomes are influenced by a range of post-event factors such as road design, vehicle safety features, emergency response time and available of medical services.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

This research investigates vehicle-pedestrian interactions in two distinct urban environments within the Kathmandu Valley—Shantinagar (urban road) and Dhobighat (arterial road). A combination of statistical modeling using Post-Encroachment Time, survival analysis, and severity analysis provided a comprehensive framework to assess both the risk and severity of these interactions. Comparative analysis of the two sites revealed notable behavioral and environmental contrasts. Shantinagar, characterized by narrow streets and a high volume of two-wheelers, exhibited more assertive pedestrian behavior and shorter PET values. In contrast, Dhobighat, marked by higher vehicle speeds and a greater presence of heavy vehicles, saw more cautious pedestrian behavior and longer PET values. Survival modeling further emphasized the differences between the two sites, with Dhobighat exhibiting a sharper reduction in collision risk as PET increases indicated by a higher shape parameter (k). This suggests that even minor improvements in PET at Dhobighat lead to more significant safety gains compared to Shantinagar. Additionally, the higher scale parameter (λ) indicates that safer interactions tend to stabilize at higher PET values, reflecting the need for greater temporal gaps in Dhobighat's faster, heavier traffic conditions. Furthermore, the severity of these interactions was analyzed by integrating both conflict and crash data, providing a comprehensive understanding of the risks associated with vehicle-pedestrian conflicts. Following are the some major conclusions drawn based on the numerical analysis.

- The Multiple linear regression modeling of Post-Encroachment Time determined the key determinants as Curb Time, Vehicle Type, Yield, Own Path, Conflict Type, Crossing Speed, Traffic Speed, and Perceived Volume.
- PET is positively related to several factors, including Curb Time, Vehicle Type (Bus-Truck and SUV-Car), Yield (Yes), Conflict Type (Maneuver Conflict and Speed-Based Conflict), and Traffic Speed. The coefficients for these parameters were found to be 0.04, 0.42, 0.20, 0.14, 0.40, and 0.24, respectively.
- PET is negatively related to Crossing Speed, Perceived Volume, and Crossing Stage (Single), with coefficients of -0.08, -0.08, and -0.10, respectively.

- Survival analysis for urban road (Shantinagar) established Post-Encroachment Time thresholds of 1.21 seconds and 2.45 seconds, categorizing interactions into High-Risk (PET < 1.21 s), Medium-Risk (1.21 s ≤ PET < 2.45 s), and Low-Risk (PET ≥ 2.45 s) zones.
- Survival analysis for arterial road (Dhobighat) determined PET thresholds of 1.71 seconds and 2.98 seconds, classifying interactions as High-Risk (PET < 1.71 s), Medium-Risk (1.71 s ≤ PET < 2.98 s), and Low-Risk (PET ≥ 2.98 s), reflecting higher vehicle speeds and the need for longer safe gaps.
- Survival analysis for combined dataset provided generalized PET thresholds of 1.34 seconds and 2.88 seconds, applicable for mixed traffic environments, defining High-Risk (PET < 1.34 s), Medium-Risk (1.34 s ≤ PET < 2.88 s), and Low-Risk (PET ≥ 2.88 s) interactions when site-specific modeling is not feasible.
- The severity analysis revealed that buses and trucks exhibited the highest risk, with 16.90% of interactions classified as Extreme Severity.
- SUV-Cars had a substantial proportion of severe interactions, with 59.22% falling into the High to Very High Severity categories.
- Two-wheelers predominantly experienced Low to Moderate Severity, accounting for 56% of interactions, reflecting their lower kinetic energy and reduced impact potential.

5.2 Recommendation

The findings of this study highlight the unique challenges of vehicle-pedestrian interactions in Nepal's mixed traffic environment. Despite the model's moderate explanatory power, the findings underscore the need for targeted interventions such as, stricter yielding laws, infrastructure upgrades and awareness campaigns for drivers and pedestrians are essential to mitigate risks and enhance safety in high-conflict areas. Furthermore, we recommend adopting context-specific PET thresholds for risk assessment in Kathmandu Valley and other areas with similar mixed traffic scenarios. Rather than relying on conventional thresholds (e.g., 1 second and 2 seconds), adopting site-specific PET thresholds provides more accurate insights into risk dynamics. For instance, on an arterial road like Dhobighat, PET thresholds of 1.71 and 2.98 seconds can be used to classify High-Risk, Medium-Risk, and Low-Risk interactions. Similarly, for an urban road like Shantinagar, thresholds of 1.21

and 2.45 seconds are more appropriate. In cases where site-specific modeling is not feasible, thresholds derived from the combined dataset—1.34 and 2.88 seconds—may serve as representative values for broader applications in mixed traffic environments. To reduce extreme severity risks from heavy vehicles like buses and trucks, stricter midblock crosswalk regulations such as enforcing lower speed limits (e.g., 20 km/h) near the crossing should be implemented. Several behavioral and contextual variables such as number of interacted vehicles, accepted gaps, and pedestrian familiarity could improve predictive accuracy but were excluded due to time constraints and survey limitations. As the study focused on peak-hour data, future research could explore off-peak periods and adopt advanced non-linear models (e.g., random forests, gradient boosting) to better capture complex interactions. Though the current multiple linear regression (MLR) approach was prioritized for its interpretability and alignment with existing surrogate safety literature

REFERENCES

Ansariyar, A. (2023). Investigating the Car-Pedestrian Conflicts Based on an Innovative Post Encroachment Time Threshold (PET) Classification. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4377745

Arun, A., Haque, M. M., Washington, S., Sayed, T., & Mannering, F. (2021). A systematic review of traffic conflict-based safety measures with a focus on application context. *Analytic Methods in Accident Research*, *32*, 100185. https://doi.org/10.1016/J.AMAR.2021.100185

Bajracharya, S., & Marsani, A. (2021). *Study on Drivers' Yielding Behavior at Pedestrian Crossings*. *10*, 2350–8906.

Bendak, S., Alnaqbi, A. M., Alzarooni, M. Y., Aljanaahi, S. M., & Alsuwaidi, S. J. (2021). Factors affecting pedestrian behaviors at signalized crosswalks: An empirical study. *Journal of Safety Research*, *76*, 269–275. https://doi.org/10.1016/j.jsr.2020.12.019

Burch, C., Cook, L., & Dischinger, P. (2014). A Comparison of KABCO and AIS Injury Severity Metrics Using CODES Linked Data. *Traffic Injury Prevention*, *15*(6), 627–630. https://doi.org/10.1080/15389588.2013.854348

Chen, Y., Ma, J., & Chen, N. (2019). Analyzing Pedestrian Fatality Risk in Accidents at Mid-Blocks. *Journal of Transportation Technologies*, *09*(02), 171–192. https://doi.org/10.4236/jtts.2019.92011

Federal Democratic Republic of Nepal Department of Roads (DOR) Kathmandu Valley Development Authority (KVDA) FEDERAL DEMOCRATIC REPUBLIC OF NEPAL THE PROJECT ON URBAN TRANSPORT IMPROVEMENT FOR KATHMANDU VALLEY FINAL REPORT VOLUME I PRESENT CONDITIONS. (2017).

FULLER, R. (1984). A conceptualization of driving behaviour as threat avoidance. *Ergonomics*, *27*(11), 1139–1155. https://doi.org/10.1080/00140138408963596

Gautam, S., & Joshi, B. (2024). Road Traffic Accident and its Characteristics in Kathmandu Valley. *International Journal on Engineering Technology*, *1*(2), 156–163. https://doi.org/10.3126/injet.v1i2.66726

Govinda, L., Sai Kiran Raju, M. R., & Ravi Shankar, K. V. R. (2022). Pedestrian-vehicle interaction severity level assessment at uncontrolled intersections using machine learning algorithms. *Safety Science*, *153*, 105806. https://doi.org/10.1016/J.SSCI.2022.105806

Harkey, D. L., Reinfurt, D. W., & Knuiman, M. (1998). Development of the Bicycle Compatibility Index. *Transportation Research Record: Journal of the Transportation Research Board*, *1636*(1), 13–20. https://doi.org/10.3141/1636-03

Johnsson, C., Laureshyn, A., & De Ceunynck, T. (2018). In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport Reviews*, *38*(6), 765–785. https://doi.org/10.1080/01441647.2018.1442888

Kadali, B. R., & Vedagiri, P. (2016). Proactive pedestrian safety evaluation at unprotected mid-block crosswalk locations under mixed traffic conditions. *Safety Science*, *89*, 94–105. https://doi.org/10.1016/j.ssci.2016.05.014

Karkee, R., & Lee, A. H. (2016). Epidemiology of road traffic injuries in Nepal, 2001–2013: systematic review and secondary data analysis. *BMJ Open*, *6*(4), e010757. https://doi.org/10.1136/bmjopen-2015-010757

Kc, A., Pradhananga, R., Koju, U., & Rojee Pradhananga, A. K. C. (2024). *Spatial Analysis of Road Traffic Crash Hotspots in Kathmandu Valley, Nepal IOE Graduate Conference Spatial Analysis of Road Traffic Crash Hotspots in Kathmandu Valley, Nepal*. https://www.researchgate.net/publication/388531875

Koh, P. P., Wong, Y. D., & Chandrasekar, P. (2014). Safety evaluation of pedestrian behaviour and violations at signalised pedestrian crossings. *Safety Science*, *70*, 143–152. https://doi.org/10.1016/J.SSCI.2014.05.010

Marisamynathan, S., & Vedagiri, P. (2020). PEDESTRIAN SAFETY EVALUATION OF SIGNALIZED INTERSECTIONS USING SURROGATE SAFETY MEASURES. *Transport*, *35*(1), 48–56. https://doi.org/10.3846/transport.2020.12157

Mohamed, M. G., & Saunier, N. (2018). The impact of motion prediction methods on surrogate safety analysis: A case study of left-turn and opposite-direction interactions at a signalized intersection in Montreal. *Journal of Transportation Safety & Security*, *10*(4), 265–287. https://doi.org/10.1080/19439962.2016.1255690

Mohan, D., Tiwari, G., & Bhalla, K. (n.d.). *T R I P P TRANSPORTATION RESEARCH AND INJURY PREVENTION PROGRAMME ROAD SAFETY IN INDIA STATUS REPORT 2015*. http://tripp.iitd.ernet.in/

Nemeth, B., & Menk, I. (2014). *Uncontrolled Pedestrian Crossing Evaluation Incorporating Highway Capacity Manual Unsignalized Pedestrian Crossing Analysis*

Methodology. http://www.mnltap.umn.edu/publications/handbooks/pedcrossingguide/

Ozili, P. K. (2022). The Acceptable R-Square in Empirical Modelling for Social Science Research. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.4128165

Peesapati, L. N., Hunter, M. P., & Rodgers, M. O. (2018). Can post encroachment time substitute intersection characteristics in crash prediction models? *Journal of Safety Research*, *66*, 205–211. https://doi.org/10.1016/J.JSR.2018.05.002

Piyalungka, S., Kanitpong, K., & Karoonsoontawong, A. (2025). Pedestrian gap acceptance behavior at unsignalized mid-block crossing under mixed traffic conditions. *IATSS Research*, *49*(2), 105–113. https://doi.org/10.1016/j.iatssr.2025.03.001

Risser, R. (1985). Behavior in traffic conflict situations. *Accident Analysis & Prevention*, *17*(2), 179–197. https://doi.org/10.1016/0001-4575(85)90020-X

Shah, D. R., & Pradhananga, R. (2024). Assessment of pedestrians' red light violation behavior at signalized crosswalks in Kathmandu, Nepal. *Transportation Research Interdisciplinary Perspectives*, *24*, 101035. https://doi.org/10.1016/j.trip.2024.101035

Shilpakar, R., & Shahi, P. B. (2022). *Proceedings of 12 th IOE Graduate Conference An assessment of Pedestrian Compliance on Using Foot Over Bridge: A Case Study of Kathmandu Valley*.

Stevenson, M., Segui-Gomez, M., Lescohier, I., Di Scala, C., & McDonald-Smith, G. (2001). An overview of the injury severity score and the new injury severity score. *Injury Prevention*, *7*(1), 10–13. https://doi.org/10.1136/ip.7.1.10

Šucha, M. (2014). Road users' strategies and communication: driver-pedestrian interaction. In *Transport Research Arena*.

Sucha, M., Dostal, D., & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention*, *102*, 41–50. https://doi.org/10.1016/j.aap.2017.02.018

Tageldin, A., Sayed, T., & Shaaban, K. (2017). Comparison of Time-Proximity and Evasive Action Conflict Measures: Case Studies from Five Cities. *Transportation Research Record: Journal of the Transportation Research Board*, *2661*(1), 19–29. https://doi.org/10.3141/2661-03

Thompson, L. L., Rivara, F. P., Ayyagari, R. C., & Ebel, B. E. (2013). Impact of social and technological distraction on pedestrian crossing behaviour: an observational study. *Injury Prevention*, *19*(4), 232–237. https://doi.org/10.1136/injuryprev-2012-040601

VAN DER HORST, A. (1990). *TIME-BASED ANALYSIS OF ROAD USE BEHAVIOR IN NORMAL AND CRITICAL ENCOUNTERS*.

Várhelyi, A. (1998). Drivers' speed behaviour at a zebra crossing: a case study. *Accident Analysis & Prevention*, *30*(6), 731–743. https://doi.org/10.1016/S0001-4575(98)00026-8

Vedagiri, P., & Kadali, B. R. (2016). Evaluation of Pedestrian–Vehicle Conflict Severity at Unprotected Midblock Crosswalks in India. *Https://Doi.Org/10.3141/2581-06*, *2581*, 48–56. https://doi.org/10.3141/2581-06

WHO. (2023). *Global Status Report on Road Safety 2023*.

Zheng, L., Sayed, T., & Essa, M. (2019). Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. *Accident Analysis & Prevention*, *123*, 314–323. https://doi.org/10.1016/J.AAP.2018.12.007

APPENDIX A : CODE in PYTHON

Multiple Linear Regression Model

```
import pandas as pd
import statsmodels.api as sm
from statsmodels.stats.outliers_influence import variance_inflation_factor
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OrdinalEncoder, OneHotEncoder, StandardScaler

# Load data
df  # assuming your DataFrame is already loaded

# Define variables
X = df.drop("PET", axis=1)
y = df["PET"]

# Preprocessing settings
ordinal_features = ['Age', 'Size', 'Curb Time', 'Exposure Time']
ordinal_categories = [['<20', '20-50', '>50'],
                      ['Size_1', 'Size_2', 'Size>2'],
                      ['NWT', 'SWT', 'LWT'],
                      ['NET', 'SET', 'LET']]

nominal_features = ['Gender', 'Own Path', 'Crossing Pattern', 'Vehicle Type', 'Yield', 'Crossing Stage', 'Conflict Location','Conflict Type']
numeric_features = ['Traffic Speed', 'Crossing Speed','Perceived Vol']

# Preprocessing pipeline
preprocessor = ColumnTransformer(
    transformers=[
        ('ordinal', OrdinalEncoder(categories=ordinal_categories), ordinal_features),
        ('nominal', OneHotEncoder(drop='first'), nominal_features),
        ('numeric', StandardScaler(), numeric_features)
    ],
    remainder='passthrough'
)
# Apply preprocessing
X_preprocessed = preprocessor.fit_transform(X)

# Convert the transformed data back to DataFrame
feature_names = preprocessor.get_feature_names_out()
X_sm = pd.DataFrame(X_preprocessed, columns=feature_names, index=X.index)

# Ensure y has the same index as X_sm
y = y.loc[X_sm.index]
```

```
# Add intercept for statsmodels
X_sm = sm.add_constant(X_sm)

# Fit the OLS model
model = sm.OLS(y, X_sm).fit()
print(model.summary())

# Calculate VIF
vif_data = pd.DataFrame()
vif_data["Feature"] = X_sm.columns
vif_data["VIF"] = [variance_inflation_factor(X_sm.values, i) for i in range(X_sm.shape[1])]
print("\nVIF Scores:")
print(vif_data)
```

SURVIVAL ANALYSIS

```
import numpy as np
import pandas as pd
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt

# Step 1: Load your data
data = pd.read_csv('CombinedfulldataWeibull.csv')
PET = data['PET Unique'].values  # Independent variable (PET)
P = data['S(t)'].values  # Dependent variable (Survival Probability)

# Step 2: Define the Weibull CDF function
def weibull_cdf(t, k, lambda_param):
    return 1 - np.exp(-((t / lambda_param) ** k))

# Step 3: Fit the Weibull CDF to the data
initial_guess = [2, 2]  # Initial guesses for k and lambda
params, _ = curve_fit(weibull_cdf, PET, P, p0=initial_guess)
k_fitted, lambda_fitted = params
print(f"Fitted Shape Parameter (k): {k_fitted:.2f}")
print(f"Fitted Scale Parameter (λ): {lambda_fitted:.2f}")

# Step 4: Define the Weibull derivative function
def weibull_derivative(t, k, lam):
    return (k / lam) * ((t / lam) ** (k - 1)) * np.exp(-((t / lam) ** k))

# Step 5: Generate PET range and calculate CDF and derivative
PET_range = np.linspace(min(PET), max(PET), 1000)  # Fine-grained PET values
cdf_fit = weibull_cdf(PET_range, k_fitted, lambda_fitted)  # Fitted CDF
slope = weibull_derivative(PET_range, k_fitted, lambda_fitted)  # Derivative of the CDF

# Step 6: Identify thresholds
threshold_1_index = np.argmax(slope)  # Peak of the derivative
threshold_1 = PET_range[threshold_1_index]
flattening_threshold = 0.01  # Define flattening threshold for derivative
flattening_index = np.where(slope < flattening_threshold)[0][0]
threshold_2 = PET_range[flattening_index]
print(f"Threshold 1 (High to Medium Risk): {threshold_1:.2f}")
print(f"Threshold 2 (Medium to Low Risk): {threshold_2:.2f}")

# Step 7: Step-by-Step Visualizations
# Plot 1: Data Points and Weibull CDF Fit
plt.figure(figsize=(10, 6))
plt.scatter(PET, P, color='blue', label='Data Points', s=50)
plt.plot(PET_range, cdf_fit, color='red', lw=2, label='Fitted Weibull CDF')
plt.title('Weibull CDF Fit to Data', fontsize=14)
plt.xlabel('PET', fontsize=12)
plt.ylabel('Survival Probability', fontsize=12)
```

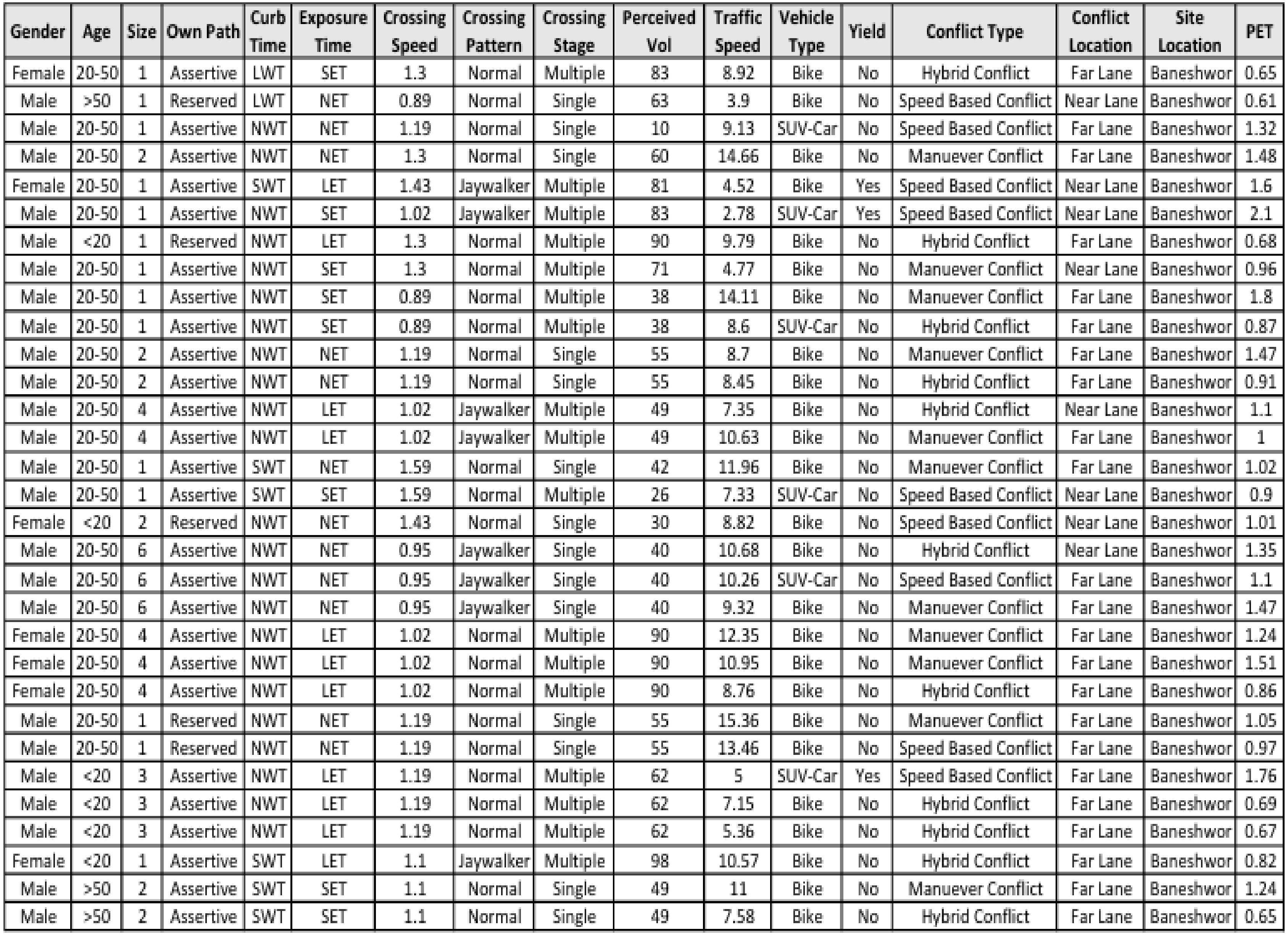
```
plt.legend(fontsize=12)
plt.grid(True)
plt.show()

# Plot 2: Rate of Change of the CDF (Derivative)
plt.figure(figsize=(10, 6))
plt.plot(PET_range, slope, label='Rate of Change (Weibull Derivative)', color='orange')
plt.axvline(x=threshold_1, color='red', linestyle='--', label=f'Threshold 1
(PET={threshold_1:.2f})')
plt.axvline(x=threshold_2, color='green', linestyle='--', label=f'Threshold 2
(PET={threshold_2:.2f})')
plt.title('Rate of Change and Risk Thresholds', fontsize=14)
plt.xlabel('PET', fontsize=12)
plt.ylabel('Rate of Change', fontsize=12)
plt.legend(fontsize=12)
plt.grid(True)
plt.show()

# Plot 3: Final Classification Curve
plt.figure(figsize=(10, 6))
plt.scatter(PET, P, color='blue', label='Data Points', s=50)
plt.plot(PET_range, cdf_fit, color='red', lw=2, label='Fitted Weibull CDF')
plt.axvline(x=threshold_1, color='red', linestyle='--', label=f'Threshold 1
(PET={threshold_1:.2f})')
plt.axvline(x=threshold_2, color='green', linestyle='--', label=f'Threshold 2
(PET={threshold_2:.2f})')
plt.fill_between(PET_range, 0, cdf_fit, where=(PET_range <= threshold_1),
color='yellow', alpha=0.3, label='High Risk Region')
plt.fill_between(PET_range, 0, cdf_fit, where=((PET_range > threshold_1) & (PET_range
<= threshold_2)), color='cyan', alpha=0.3, label='Medium Risk Region')
plt.fill_between(PET_range, 0, cdf_fit, where=(PET_range > threshold_2), color='purple',
alpha=0.3, label='Low Risk Region')
plt.title('Final Classification of Risk Regions', fontsize=14)
plt.xlabel('PET', fontsize=12)
plt.ylabel('Survival Probability', fontsize=12)
plt.legend(fontsize=12)
plt.grid(True)
plt.show()

# Step 8: Categorize PET values into risk levels
def categorize_risk(pet_value):
    if pet_value < threshold_1:
        return 'High Risk'
    elif pet_value < threshold_2:
        return 'Medium Risk'
    else:
        return 'Low Risk'

data['Risk Category'] = data['PET Unique'].apply(categorize_risk)
```

```
# Save the updated dataset to a CSV file
data.to_csv('Updated_Data_with_Risk_Categories.csv', index=False)
print("Updated dataset saved as 'Updated_Data_with_Risk_Categories.csv'.")

# Display the updated dataset with risk categories
#print(data[['PET_x', 'Risk Category']])
```

SEVERITY ANALYSIS

```
import numpy as np
import pandas as pd

# Compute the 5th, 25th, 50th, 75th, and 95th percentiles
severity_percentiles = df_cleaned["SeverityIndex"].quantile([0.05, 0.25, 0.50, 0.75, 0.95])
p5, p25, p50, p75, p95 = severity_percentiles.values

# Print percentiles
print(f"5th Percentile: {p5}\n25th Percentile: {p25}\n50th Percentile (Median): {p50}\n75th Percentile: {p75}\n95th Percentile: {p95}")

# Define severity categorization function
def categorize_severity(value):
    if value <= p5:
        return "Very Low"
    elif value <= p25:
        return "Low"
    elif value <= p50:
        return "Moderate"
    elif value <= p75:
        return "High"
    elif value <= p95:
        return "Very High"
    else:
        return "Extreme"

# Apply categorization
df_cleaned["Severity Level"] = df_cleaned["SeverityIndex"].apply(categorize_severity)

# Define the correct order of severity levels
severity_order = ["Very Low", "Low", "Moderate", "High", "Very High", "Extreme"]

# Compute percentage distribution of severity levels for each vehicle type
severity_distribution = df_cleaned.groupby("Vehicle Type")["Severity Level"].value_counts(normalize=True) * 100
severity_distribution = severity_distribution.unstack()

# Reorder columns based on severity levels
severity_distribution = severity_distribution[severity_order]

# Print the severity distribution table
print("🔹 Percentage Distribution of Severity Levels for Each Vehicle Type:")
print(severity_distribution)
```

APPENDIX B : SAMPLE DATA



Gender	Age	Size	Own Path	Curb Time	Exposure Time	Crossing Speed	Crossing Pattern	Crossing Stage	Perceived Vol	Traffic Speed	Vehicle Type	Yield	Conflict Type	Conflict Location	Site Location	PET
Female	20-50	1	Assertive	LWT	SET	1.3	Normal	Multiple	83	8.92	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.65
Male	>50	1	Reserved	LWT	NET	0.89	Normal	Single	63	3.9	Bike	No	Speed Based Conflict	Near Lane	Baneshwor	0.61
Male	20-50	1	Assertive	NWT	NET	1.19	Normal	Single	10	9.13	SUV-Car	No	Speed Based Conflict	Far Lane	Baneshwor	1.32
Male	20-50	2	Assertive	NWT	NET	1.3	Normal	Single	60	14.66	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.48
Female	20-50	1	Assertive	SWT	LET	1.43	Jaywalker	Multiple	81	4.52	Bike	Yes	Speed Based Conflict	Near Lane	Baneshwor	1.6
Male	20-50	1	Assertive	NWT	SET	1.02	Jaywalker	Multiple	83	2.78	SUV-Car	Yes	Speed Based Conflict	Near Lane	Baneshwor	2.1
Male	<20	1	Reserved	NWT	LET	1.3	Normal	Multiple	90	9.79	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.68
Male	20-50	1	Assertive	NWT	SET	1.3	Normal	Multiple	71	4.77	Bike	No	Manuever Conflict	Near Lane	Baneshwor	0.96
Male	20-50	1	Assertive	NWT	SET	0.89	Normal	Multiple	38	14.11	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.8
Male	20-50	1	Assertive	NWT	SET	0.89	Normal	Multiple	38	8.6	SUV-Car	No	Hybrid Conflict	Far Lane	Baneshwor	0.87
Male	20-50	2	Assertive	NWT	NET	1.19	Normal	Single	55	8.7	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.47
Male	20-50	2	Assertive	NWT	NET	1.19	Normal	Single	55	8.45	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.91
Male	20-50	4	Assertive	NWT	LET	1.02	Jaywalker	Multiple	49	7.35	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	1.1
Male	20-50	4	Assertive	NWT	LET	1.02	Jaywalker	Multiple	49	10.63	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1
Male	20-50	1	Assertive	SWT	NET	1.59	Normal	Single	42	11.96	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.02
Male	20-50	1	Assertive	SWT	SET	1.59	Normal	Multiple	26	7.33	SUV-Car	No	Speed Based Conflict	Near Lane	Baneshwor	0.9
Female	<20	2	Reserved	NWT	NET	1.43	Normal	Single	30	8.82	Bike	No	Speed Based Conflict	Near Lane	Baneshwor	1.01
Male	20-50	6	Assertive	NWT	NET	0.95	Jaywalker	Single	40	10.68	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	1.35
Male	20-50	6	Assertive	NWT	NET	0.95	Jaywalker	Single	40	10.26	SUV-Car	No	Speed Based Conflict	Far Lane	Baneshwor	1.1
Male	20-50	6	Assertive	NWT	NET	0.95	Jaywalker	Single	40	9.32	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.47
Female	20-50	4	Assertive	NWT	LET	1.02	Normal	Multiple	90	12.35	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.24
Female	20-50	4	Assertive	NWT	LET	1.02	Normal	Multiple	90	10.95	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.51
Female	20-50	4	Assertive	NWT	LET	1.02	Normal	Multiple	90	8.76	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.86
Male	20-50	1	Reserved	NWT	NET	1.19	Normal	Single	55	15.36	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.05
Male	20-50	1	Reserved	NWT	NET	1.19	Normal	Single	55	13.46	Bike	No	Speed Based Conflict	Far Lane	Baneshwor	0.97
Male	<20	3	Assertive	NWT	LET	1.19	Normal	Multiple	62	5	SUV-Car	Yes	Speed Based Conflict	Far Lane	Baneshwor	1.76
Male	<20	3	Assertive	NWT	LET	1.19	Normal	Multiple	62	7.15	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.69
Male	<20	3	Assertive	NWT	LET	1.19	Normal	Multiple	62	5.36	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.67
Female	<20	1	Assertive	SWT	LET	1.1	Jaywalker	Multiple	98	10.57	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.82
Male	>50	2	Assertive	SWT	SET	1.1	Normal	Single	49	11	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.24
Male	>50	2	Assertive	SWT	SET	1.1	Normal	Single	49	7.58	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.65

Male	20-50	3	Reserved	NWT	LET	1.43	Normal	Multiple	55	7.75	Bus-Truck	No	Speed Based Conflict	Far Lane	Baneshwor	1.5
Male	20-50	1	Assertive	NWT	NET	1.19	Normal	Single	20	10.83	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	1.03
Female	<20	1	Assertive	NWT	NET	1.59	Normal	Single	27	6.96	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	0.58
Female	20-50	1	Reserved	NWT	NET	1.19	Normal	Single	35	8.34	Bike	No	Speed Based Conflict	Near Lane	Baneshwor	0.87
Female	20-50	1	Reserved	NWT	NET	1.19	Normal	Single	35	9.36	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.9
Female	20-50	1	Reserved	NWT	NET	1.3	Normal	Single	11	11.24	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.54
Male	>50	1	Reserved	NWT	NET	0.95	Normal	Single	44	8.3	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	1.3
Male	>50	1	Reserved	LWT	LET	1.19	Jaywalker	Single	79	7.6	Bike	No	Speed Based Conflict	Far Lane	Baneshwor	0.65
Female	<20	1	Reserved	LWT	NET	1.19	Normal	Single	118	6.61	SUV-Car	No	Speed Based Conflict	Far Lane	Baneshwor	1.01
Male	20-50	3	Assertive	NWT	NET	1.14	Normal	Single	34	14.19	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.21
Female	<20	2	Assertive	NWT	LET	1.3	Normal	Multiple	137	10	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	0.76
Female	<20	3	Assertive	SWT	LET	1.02	Normal	Multiple	116	11.41	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.01
Female	<20	3	Assertive	SWT	LET	1.02	Normal	Multiple	116	8.52	SUV-Car	No	Speed Based Conflict	Far Lane	Baneshwor	0.76
Male	20-50	3	Assertive	NWT	SET	1.19	Normal	Single	50	11.29	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.34
Female	<20	2	Reserved	LWT	SET	1.02	Normal	Multiple	56	10.49	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	1.23
Female	20-50	2	Reserved	LWT	SET	1.19	Normal	Single	28	5.11	Bike	No	Speed Based Conflict	Near Lane	Baneshwor	0.72
Male	20-50	1	Reserved	NWT	NET	1.3	Normal	Single	44	8.62	SUV-Car	No	Speed Based Conflict	Far Lane	Baneshwor	0.69
Male	20-50	3	Assertive	NWT	NET	1.19	Normal	Single	55	5.56	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	0.51
Male	20-50	3	Assertive	NWT	NET	1.19	Normal	Single	55	10.6	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.63
Female	20-50	1	Reserved	NWT	NET	1.02	Normal	Single	69	9.8	Bus-Truck	No	Speed Based Conflict	Far Lane	Baneshwor	1.05
Female	20-50	1	Reserved	NWT	NET	1.02	Normal	Single	69	6.55	Bike	Yes	Speed Based Conflict	Far Lane	Baneshwor	1.16
Male	20-50	2	Reserved	NWT	NET	0.99	Jaywalker	Single	108	9.96	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.07
Female	20-50	1	Reserved	SWT	NET	1.02	Normal	Single	92	9.24	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	0.77
Male	20-50	1	Assertive	NWT	NET	1.3	Jaywalker	Single	55	7.36	Bike	No	Speed Based Conflict	Far Lane	Baneshwor	0.6
Male	<20	1	Reserved	NWT	NET	1.19	Normal	Single	55	12.35	Bike	No	Manuever Conflict	Far Lane	Baneshwor	0.99
Male	20-50	2	Reserved	SWT	NET	1.3	Normal	Single	26	10.03	Bike	No	Speed Based Conflict	Far Lane	Baneshwor	1.51
Male	20-50	2	Reserved	SWT	NET	1.3	Normal	Single	26	9.33	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	1.07
Female	20-50	2	Assertive	LWT	SET	1.19	Normal	Single	67	7.77	Bike	No	Hybrid Conflict	Far Lane	Baneshwor	1.01
Female	20-50	3	Assertive	SWT	LET	1.02	Normal	Single	67	10.03	Bike	No	Hybrid Conflict	Near Lane	Baneshwor	0.87
Female	20-50	3	Assertive	SWT	LET	1.02	Normal	Single	67	3.84	Bike	No	Speed Based Conflict	Far Lane	Baneshwor	1.42
Male	20-50	5	Assertive	SWT	LET	1.3	Normal	Single	82	9.22	Bike	No	Manuever Conflict	Near Lane	Baneshwor	1.16
Male	20-50	1	Reserved	SWT	NET	1.43	Normal	Single	53	11.7	Bike	No	Manuever Conflict	Near Lane	Baneshwor	1.24
Male	20-50	1	Reserved	LWT	SET	1.19	Normal	Single	84	7.62	Bike	No	Manuever Conflict	Far Lane	Baneshwor	1.73

Female	20-50	1	Assertive	NWT	SET	1.36	Normal	Multiple	51	5.72	Bike	Yes	Hybrid Conflict	Far Lane	Dhobighat	1.3
Female	20-50	1	Reserved	LWT	NET	1.39	Jaywalker	Single	62	7.77	Bus-Truck	No	Speed Based Conflict	Near Lane	Dhobighat	2.81
Female	20-50	1	Reserved	NWT	NET	1.61	Jaywalker	Single	18	9.89	Bus-Truck	No	Hybrid Conflict	Near Lane	Dhobighat	1.06
Male	>50	3	Reserved	NWT	NET	1.14	Normal	Single	4	10.4	Bike	No	Hybrid Conflict	Near Lane	Dhobighat	1.21
Female	20-50	2	Reserved	NWT	SET	1.27	Jaywalker	Single	24	18.61	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.96
Male	20-50	1	Reserved	NWT	SET	1.01	Normal	Multiple	63	13.64	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.45
Female	20-50	1	Reserved	NWT	NET	1.3	Normal	Single	20	8.9	SUV-Car	Yes	Speed Based Conflict	Far Lane	Dhobighat	2.26
Female	20-50	1	Reserved	NWT	NET	1.3	Normal	Single	20	10.06	Bike	No	Hybrid Conflict	Far Lane	Dhobighat	1.1
Female	20-50	3	Reserved	NWT	LET	0.9	Normal	Single	15	4.87	SUV-Car	Yes	Speed Based Conflict	Near Lane	Dhobighat	2.29
Female	20-50	3	Reserved	NWT	LET	0.9	Normal	Single	15	12.62	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.62
Female	<20	1	Reserved	LWT	SET	0.71	Normal	Single	50	10.67	SUV-Car	No	Manuever Conflict	Far Lane	Dhobighat	2.2
Female	<20	1	Reserved	NWT	NET	1.14	Normal	Single	13	10.53	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.43
Female	20-50	1	Reserved	NWT	SET	1.09	Normal	Multiple	16	7.84	SUV-Car	Yes	Speed Based Conflict	Far Lane	Dhobighat	1.45
Male	20-50	1	Reserved	NWT	SET	1.29	Normal	Multiple	41	8.44	SUV-Car	No	Manuever Conflict	Far Lane	Dhobighat	2.21
Male	20-50	1	Reserved	NWT	SET	1.29	Normal	Multiple	41	7.6	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.85
Female	>50	1	Reserved	NWT	LET	1.17	Normal	Multiple	39	8.67	Bike	No	Manuever Conflict	Far Lane	Dhobighat	2.11
Male	20-50	2	Assertive	LWT	NET	1.25	Normal	Single	78	5.04	SUV-Car	Yes	Speed Based Conflict	Far Lane	Dhobighat	1.8
Female	20-50	2	Reserved	LWT	SET	1.18	Normal	Multiple	62	10.6	Bus-Truck	No	Speed Based Conflict	Near Lane	Dhobighat	1.96
Male	20-50	1	Reserved	NWT	NET	1.66	Jaywalker	Single	38	7.68	SUV-Car	Yes	Speed Based Conflict	Near Lane	Dhobighat	1.56
Male	20-50	1	Reserved	NWT	NET	1.14	Normal	Single	30	9.69	SUV-Car	No	Manuever Conflict	Near Lane	Dhobighat	2.1
Female	>50	1	Reserved	NWT	SET	1.05	Normal	Multiple	31	8.92	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.94
Male	20-50	1	Reserved	NWT	SET	1.66	Jaywalker	Single	54	7.21	SUV-Car	No	Speed Based Conflict	Near Lane	Dhobighat	2.22
Female	20-50	3	Reserved	NWT	SET	1.23	Jaywalker	Single	25	12.48	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.5
Male	20-50	1	Reserved	NWT	LET	1.43	Normal	Single	39	7.38	SUV-Car	No	Manuever Conflict	Near Lane	Dhobighat	2.23
Male	>50	1	Reserved	LWT	NET	1.02	Normal	Single	57	9.97	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.68
Male	<20	1	Reserved	NWT	LET	1.64	Jaywalker	Single	27	9.07	Bike	Yes	Speed Based Conflict	Far Lane	Dhobighat	1.52
Male	20-50	1	Reserved	NWT	SET	1.35	Normal	Multiple	41	11.01	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	1.35
Female	20-50	1	Reserved	LWT	NET	1.21	Normal	Single	55	12.29	Bike	No	Manuever Conflict	Near Lane	Dhobighat	1.99
Female	20-50	1	Reserved	NWT	NET	1.13	Normal	Single	17	11.68	SUV-Car	No	Manuever Conflict	Far Lane	Dhobighat	2.04
Female	20-50	1	Reserved	NWT	NET	1.13	Normal	Single	17	9.06	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.51
Female	20-50	2	Reserved	NWT	NET	1.08	Normal	Single	12	16	Bike	No	Manuever Conflict	Near Lane	Dhobighat	1.98
Male	20-50	1	Reserved	NWT	NET	0.96	Normal	Single	40	10.41	Bike	No	Speed Based Conflict	Near Lane	Dhobighat	2.1
Male	>50	1	Reserved	NWT	LET	1.49	Normal	Multiple	63	10.16	Bus-Truck	No	Speed Based Conflict	Far Lane	Dhobighat	2.66

Male	>50	1	Reserved	NWT	NET	1.23	Normal	Single	33	12.37	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.72
Female	<20	1	Reserved	NWT	NET	1.68	Normal	Single	32	6.73	Bike	No	Hybrid Conflict	Far Lane	Dhobighat	1.64
Female	<20	1	Reserved	NWT	NET	1.68	Normal	Single	32	7.26	Bike	No	Hybrid Conflict	Near Lane	Dhobighat	1.35
Male	<20	1	Reserved	NWT	NET	1.03	Normal	Single	8	8.33	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	1.05
Female	>50	1	Assertive	NWT	SET	0.97	Normal	Multiple	82	11.78	Bike	Yes	Manuever Conflict	Far Lane	Dhobighat	1.89
Male	20-50	2	Reserved	NWT	NET	1.03	Normal	Single	16	12.46	Bike	No	Manuever Conflict	Far Lane	Dhobighat	2.11
Female	20-50	1	Reserved	NWT	SET	0.9	Normal	Multiple	44	10.33	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.73
Female	20-50	1	Reserved	NWT	SET	0.9	Normal	Multiple	44	9.95	Bus-Truck	No	Speed Based Conflict	Far Lane	Dhobighat	2.8
Female	20-50	1	Reserved	NWT	SET	0.9	Normal	Multiple	44	9.32	SUV-Car	Yes	Hybrid Conflict	Far Lane	Dhobighat	1.44
Female	20-50	2	Assertive	NWT	SET	1.05	Normal	Multiple	12	5.41	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	0.99
Male	>50	1	Reserved	NWT	LET	1.1	Normal	Multiple	65	8.4	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	1.9
Female	20-50	1	Reserved	NWT	NET	1.32	Normal	Single	5	7.04	SUV-Car	No	Speed Based Conflict	Near Lane	Dhobighat	2.13
Male	<20	1	Reserved	NWT	LET	1.22	Jaywalker	Multiple	79	7.31	Bike	No	Speed Based Conflict	Near Lane	Dhobighat	1.77
Female	<20	1	Reserved	NWT	NET	1.31	Normal	Single	20	8	Bus-Truck	No	Speed Based Conflict	Near Lane	Dhobighat	1.6
Male	20-50	1	Reserved	SWT	SET	0.79	Jaywalker	Multiple	19	12.56	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.79
Female	<20	1	Reserved	LWT	NET	1.52	Normal	Single	54	14.91	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.72
Male	20-50	1	Assertive	NWT	NET	1.52	Normal	Single	23	11.15	Bike	No	Speed Based Conflict	Near Lane	Dhobighat	1.37
Male	<20	1	Reserved	NWT	NET	1.17	Normal	Single	4	8.29	Bus-Truck	No	Speed Based Conflict	Far Lane	Dhobighat	1.96
Male	<20	1	Assertive	NWT	NET	1.26	Normal	Single	19	12.86	SUV-Car	No	Manuever Conflict	Near Lane	Dhobighat	1.76
Male	20-50	2	Reserved	NWT	NET	1.07	Jaywalker	Single	20	11.57	Bike	No	Manuever Conflict	Far Lane	Dhobighat	2.13
Female	<20	2	Assertive	NWT	LET	0.99	Normal	Multiple	69	7.12	SUV-Car	No	Hybrid Conflict	Far Lane	Dhobighat	1.64
Female	<20	2	Assertive	NWT	LET	0.99	Normal	Multiple	69	4.96	SUV-Car	Yes	Speed Based Conflict	Far Lane	Dhobighat	1.67
Female	20-50	1	Assertive	NWT	LET	1.38	Normal	Multiple	60	11.59	SUV-Car	No	Speed Based Conflict	Near Lane	Dhobighat	1.64
Female	20-50	1	Assertive	NWT	LET	1.38	Normal	Multiple	60	5.37	SUV-Car	Yes	Speed Based Conflict	Far Lane	Dhobighat	1.96
Female	>50	2	Reserved	NWT	SET	1.09	Normal	Multiple	21	9.99	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.13
Male	<20	1	Reserved	NWT	SET	1.26	Normal	Multiple	35	10.56	SUV-Car	No	Speed Based Conflict	Far Lane	Dhobighat	1.76
Female	20-50	3	Assertive	NWT	SET	1.02	Normal	Multiple	44	7.43	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.43
Male	<20	1	Reserved	LWT	NET	1.19	Normal	Single	43	10.47	Bike	No	Manuever Conflict	Far Lane	Dhobighat	1.7
Male	20-50	3	Reserved	NWT	SET	1.04	Normal	Multiple	13	6.88	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	2.19
Female	20-50	3	Reserved	NWT	NET	1.23	Normal	Single	23	10.95	Bike	Yes	Manuever Conflict	Far Lane	Dhobighat	2.43
Female	<20	1	Reserved	NWT	NET	1.12	Normal	Single	34	6.25	Bus-Truck	No	Hybrid Conflict	Far Lane	Dhobighat	1.39
Female	<20	1	Reserved	NWT	NET	1.12	Normal	Single	34	5.6	Bike	No	Speed Based Conflict	Far Lane	Dhobighat	0.88
Male	20-50	1	Reserved	NWT	NET	0.79	Normal	Single	93	10.51	SUV-Car	No	Speed Based Conflict	Near Lane	Dhobighat	1.85

PAPER NAME

Post Encroachment Time-Based Behavioral Analysis of Vehicle-Pedestrian Interactions at Unsignalized

AUTHOR

Sandesh Lamsal

WORD COUNT

17300 Words

CHARACTER COUNT

104249 Characters

PAGE COUNT

80 Pages

FILE SIZE

5.6MB

SUBMISSION DATE

Apr 18, 2025 10:50 PM GMT+5:45

REPORT DATE

Apr 18, 2025 10:51 PM GMT+5:45

● 5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 4% Internet database
- 3% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

● Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 10 words)
- Manually excluded text blocks

IOE Graduate Conference
[Placeholder for
Publication
Information]

Behavioral Dynamics of Vehicle Pedestrian Interaction at Unsignalized crosswalk - A PET based Study of Shantinagar-Aloknagar Crosswalk

Sandesh Lamsal [a], Rojee Pradhananga [b]

[a,b] *Department of Civil Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
✉ [a] 079mstre020.sandesh@pcampus.edu.np, [b] rojee.pradhananga@pcampus.edu.np

Abstract

This study investigates pedestrian-vehicle interactions through the analysis of Post-Encroachment Time (PET), a critical surrogate safety measure, in a high-conflict urban area. Using a combination of statistical modeling and behavioral analysis, the research identifies key factors influencing PET, including vehicle type, pedestrian behavior, and conflict dynamics. The findings reveal that larger vehicles (e.g., buses, trucks) and SUVs-Cars are associated with longer PET due to their slower maneuverability, while two-wheelers exhibit the shortest PET, reflecting higher risk due to their agility. Younger pedestrians (<20 years) face the highest risk, with the lowest median PET 1.05 seconds, likely due to impulsive crossing behavior. Assertive pedestrian behavior (61.41%) and multi-stage crossings (27.39%) further increase exposure to risk, particularly in Aloknagar Side (75.69% of conflicts). Hybrid conflicts, involving simultaneous speed and path adjustments, pose the highest risk with PET 1.00 seconds, while maneuver conflicts are relatively safer with PET 1.60 seconds. A multiple linear regression model, validated through statistical tests, explains 22.5% of the variance in PET ($R^2 = 0.225$), which is consistent with previous studies on behavioral interactions, where low R^2 values are common due to the stochastic nature of human behavior. Despite its moderate explanatory power, the model identifies key predictors, such as vehicle type and conflict dynamics, that align with theoretical expectations and provide actionable insights for traffic safety analysis.

Keywords

Vehicle-Pedestrian Interaction, Post-Encroachment Time, Pedestrian Behavior, Multiple Linear Regression

1. Introduction

1.1 Background

Road traffic injuries and fatalities remain a significant global health and development challenge, impacting millions of lives every year. According to the World Health Organization road traffic crashes have emerged as the leading cause of death for children and young people aged 5 to 29, and they rank as the 12th leading cause of death across all age groups worldwide [1]. Road user fatalities are distributed among different groups: 4-wheel vehicle occupants account for 30 % of global road fatalities, followed by pedestrians at 23 %, and powered two- and three-wheeler users at 21% [1]. This data highlights the vulnerability of pedestrians and the critical need for road safety interventions that consider pedestrian dynamics.

In Nepal, the growing rate of road traffic accidents (RTAs) underscores a similar crisis at the national level. Between 2001 and 2013, the Nepal Traffic Police recorded 95,902 crashes, resulting in 100,499 injuries and 14,512 fatalities. Alarmingly, the traffic death rate in Nepal escalated from 4 per 100,000 people in 2001–2002 to 7 per 100,000 in 2011–2012 [2]. These statistics reveal that motorcyclists and pedestrians are disproportionately affected, with higher incidents among men and people aged between 20 and 40. Factors such as risky pedestrian behavior, alcohol consumption, and substandard bus driving practices have been identified as potential contributors to this high accident rate. Kathmandu, Nepal's capital, faces critical road safety challenges as rapid urbanization drives population growth and transportation demand. Limited public transportation options have led many residents to rely on private vehicles, intensifying road congestion and encouraging faster driving, which raises risks for both drivers and pedestrians.

Additionally, inadequate infrastructure such as poorly maintained footpaths, limited crosswalks, and scarce traffic signals, especially at unsignalized intersections compounds safety issues. Frequent, unregulated vehicle-pedestrian interactions in these spaces, along with limited enforcement of traffic rules, create a hazardous environment where pedestrians and vehicles compete for road space. Given these significant risks, this study aims to examine the behavioral dynamics of vehicle-pedestrian interactions in Kathmandu, focusing on unsignalized crosswalks. Insights gained may support urban planning and policy initiatives to improve road safety and streamline interactions in Kathmandu's shared road spaces.

1.2 Research Objectives

The primary goal of this research is to analyze the dynamics of vehicle-pedestrian interactions at unsignalized crosswalks in Kathmandu. The other considered objectives are:

- To identify and analyze the factors influencing vehicle-pedestrian interactions at crosswalks
- To measure and model Post Encroachment Time (PET) as a key indicator of road user vulnerability in vehicle-pedestrian interactions.

2. Literature Review

2.1 Pedestrian Behavior

Pedestrian behavior is complex and unpredictable as they stop, gets distracted and change direction unexpectedly. Pedestrian behavior is shaped by dynamic factors like vehicle speed and traffic volumes, social factors such as upbringing, education, and negligence, and environmental factors including site conditions, visibility, and road layout, all of which influence decision-making and safety at crossings. Previous studies have investigated pedestrian behaviors during road crossings, often focusing on specific aspects [3]. For instance, study [4] analyzed rule violations at signalized crosswalks, while research in [5] examined the impact of social and technological distractions on crossing behavior. This emphasizes how crucial it is to take a comprehensive approach that takes into account a greater range of pedestrian behaviors in order to better understand their trends and enhance crossing safety measures.

2.2 Drivers Behavior

Driver's willingness to give the right of way to pedestrians at pedestrian crossing is low. Furthermore, drivers did not sufficiently lower their speeds to maintain a readiness to react to an unexpected dangerous situation. Moreover, drivers were willing to slow or stop when the speed of their vehicles was low. Drivers rarely yield to pedestrians at zebra crossings, with giving way occurring in only 5 % of cases. However, when pedestrians act assertively to force vehicles to stop, they manage to pass in nearly half of such conflict situations [6]. In Nepal, key factors influencing driver yielding behavior includes the number of pedestrians crossing, pedestrian gender, age, and speed, vehicle type and speed, driver gender and age, presence of a median strip, and road marking condition [7].

2.3 Vehicle-Pedestrian Interaction

When pedestrians and drivers attempt to occupy the same space simultaneously, it creates a high-risk situation. Both parties may expect the other to yield, leading to hesitation, confusion, or near-collisions. This struggle for space is particularly common at unsignalized crossings, where clear priority is lacking, and can result in unsafe maneuvers, as drivers may brake abruptly while pedestrians pause or dart across. This concept aligns with Fuller's Threat Avoidance Model, which explains how drivers respond to perceived threats, including interactions with pedestrians. According to [8] drivers continuously assess risks posed by others on the road and adjust their behavior based on threat perception and response capacity. When pedestrians and drivers attempt to occupy the same space, drivers make real-time threat assessments: if a pedestrian appears to yield, the driver may proceed without slowing down; if not, the perceived threat increases, prompting evasive actions. Fuller's model also highlights drivers' limitations in reacting to sudden pedestrian moves, as they may not fully prepare to yield until the threat becomes immediate, sometimes leaving insufficient time to stop. The interaction between pedestrian and vehicles occurs through different medium. Pedestrians often rely on eye contact (through glances or stares), hand waves, smiles, or nods. Drivers, meanwhile, communicate by flashing their lights, waving, or making eye contact [9]. However, for this communication to be effective, it must be clear and properly understood by both parties. Misinterpretations or lack of communication can lead to confusion, hesitation, or risky behavior, increasing the likelihood of accidents. In a survey, it was revealed that 47 % of the incidents occurred without communication, 11 % were due to a lack of necessary communication, and 42 % happened during actual communication [10]. The Vehicle Pedestrian interaction is complex because Pedestrians and drivers consider and deal with the same situations, but from the very different perspective [11]. The critical role of communication in these interactions makes essential to study not only the dynamics of vehicle-pedestrian communication but also the associated conflict measures to ensure safe interactions.

2.4 Post Encroachment Time

Post Encroachment Time (PET) is the time difference between the moment when one road user (e.g., a pedestrian or a vehicle) leaves the conflict point and when another road user enters the same conflict point. Johnsson et al. reviewed at a variety of approaches to evaluating conflicts involving vulnerable road users and came to the conclusion that no one indication is always appropriate in every situation [12]. Arun et al.conducted a systematic review of studies from the past decade (2010–2019) focusing on conflict measures. Despite the introduction of numerous conflict measures over time, the earliest metrics, such as time-to-collision and post-encroachment time, remain the most widely used, likely due to their simplicity and ease of interpretation [13] . Zheng et al. showed that combining time-to-collision with post-encroachment time enhances safety predictions, as each metric captures distinct dimensions of a conflict event [14]. Time-to-collision reflects the anticipated risk during a conflict event, whereas post-encroachment time measures the actual safety margin after the vehicles have passed one another and the conflict has ended. Tageldin et al. highlighted that the choice of conflict measures depends on the traffic environment being studied. This review categorizes traffic systems into two types: organized systems, typical in countries like the United States and the United Kingdom, and less-structured systems, common in nations such as China, India, and South Korea [15]. Since India and Nepal share similar traffic characteristics, conflict measures suitable for India are also applicable to Nepal. Time-to-collision calculations depend on precise measurements of speed and distance, which are easier in organized systems with automated tools whereas post-encroachment time requires direct observation, manageable even in less-organized environments [13]. Time-to-collision is commonly used to identify rear-end conflicts, while post-encroachment time is often used as a measure once trajectories intersect [16]. Models that integrate PET with traffic volume characteristics exhibit superior predictive capability compared to those that rely solely on PET [17]. Based on the insights from the literature review, PET along with relevant traffic metrics appears to be the most suitable conflict measure for the context of my study.

Figure 1: Shantinagar-Aloknagar Crosswalk

3. Methodology

3.1 Study Area Description

The Shantinagar-Aloknagar crosswalk as shown in Fig 1 was selected as the study area to analyze vehicle-pedestrian interactions and conflicts. A pilot survey was conducted for 30 minutes during peak hours and 30 minutes during off-peak hours to assess traffic conditions and the feasibility of data collection. The survey aimed to evaluate whether distant traffic signals caused congestion during peak hours and impacted vehicle speed or pedestrian movement. The findings indicated that back queuing from the traffic signals did not significantly disrupt vehicle flow speed near the crosswalk during peak hours, nor did it notably impede pedestrian crossings. Additionally, conflict occurrences were more frequent during peak hours. Since previous studies have also focused on peak hours due to higher vehicle speeds and interaction rates, and as the pilot study confirmed increased conflicts during this period, peak hours were chosen for the study. Also, due to side friction caused by the frequent stoppage of public transport vehicles in the service lanes, only the mid-section of the road was considered in study.

3.2 Data Collection and Extraction

Data was collected through a videographic survey from a suitable height using a GoPro Hero 11 (5.3K Linear, 30 fps, H.265). The footage was analyzed using MPC-HC tools, and data was precisely extracted frame by frame into an Excel sheet. The survey was conducted for 2 hours during peak hours over 3 weekdays, yielding approximately 600 interaction data sets that resulted into conflict were taken for analysis.

3.3 Variables Description

Various variables related to pedestrian behavior, driver behavior, and traffic stream have been considered for the study of PET. A detailed explanation of these variables is provided in Table 1.

3.4 Multiple Linear Regression Model

Multiple Linear Regression (MLR) is a statistical analysis method used to model the relationship between a continuous dependent variable and one or more explanatory variables. In this research, Post Encroachment Time (PET) is taken as the continuous dependent variable, representing the time difference between the moment one road user leaves the conflict zone and the moment another road user enters it. PET is influenced by various factors, such as pedestrian demographics, traffic characteristics, types of conflicts, and vehicle type, etc. which are treated as explanatory variables in the model. The MLR model is particularly useful for understanding how these explanatory variables collectively influence PET, providing insights into pedestrian-vehicle interactions and road safety. The basic equation of the MLR model is expressed as:

$$PET = \beta_a + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{1}$$

where, PET = Post Encroachment Time; $X_i - n$= explanatory variables;

4. Results and Discussions

4.1 Descriptive Analysis

Approximately 82.27% of vehicles do not yield to pedestrians while crossing, highlighting a significant risk to pedestrian safety. On the pedestrian side, 61.41% exhibit assertive behavior, signaling their presence on the road. The majority of conflicts (75.69%) occur at Aloknagar side, indicating a need for targeted safety measures in that lane. Additionally, 29.39% of pedestrians cross the road in multiple stages, pausing midway to wait for traffic to clear, which increases their exposure to risk. Vehicle type and volume play a crucial role in these interactions. Two-wheelers account for 68.65 % of conflicts, primarily due to their high volume in traffic and lower perceived threat resulting from their smaller vehicle size. SUVs and cars follow with 28.28 %, while conflicts involving buses and trucks (3.05%) are rare, as these vehicles primarily use service lanes for passenger pick-up and drop-off and are restricted during daytime inside the ring road Interestingly, only 18.80% of pedestrians are jaywalkers, likely due to the presence of medians with planted trees, which limit crossing points. Pedestrians are often compelled to use zebra crossings, as median openings are only available at these locations.

Table 1: Description of Variables Used in the Study

Variable	**Description**	**Type of Variable**	**Unit/Code**
Post Encroachment Time (PET)	Time difference between the moment one road user (e.g., a pedestrian) leaves the conflict zone and the moment another road user (e.g., a vehicle) enters it.	Continuous	Seconds (s)
Crossing Speed	The speed at which a pedestrian crosses the road.	Continuous	Meter per second (m/s)
Traffic Speed	The average speed of the interacting vehicle.	Continuous	Meter per second (m/s)
Perceived Volume	The number of vehicles perceived while waiting and crossing the road in each direction.	Continuous	Traffic volume per min
Gender	Based on visual appearance.	Categorical	[Male, Female]
Age	Based on visual appearance.	Categorical	[<20, 20-50, >50]
Size	Number of pedestrians crossing simultaneously at a time.	Categorical	[Size_1, Size_2, Size_>2]
Own Path	Pedestrian behavior: Assertive (standing below the curb or using hand gestures) or Reserved (standing at the curb, no gestures).	Categorical	[Assertive, Reserved]
Crossing Pattern	Pedestrian behavior: Normal (crossing correctly) or Jaywalker (crossing improperly).	Categorical	[Jaywalker, Normal]
Vehicle Type	Type of vehicle involved in the interaction (visual observation).	Categorical	[2W, SUV-Car, Bus-Truck]
Yield	Indicates whether the vehicle nearly stopped during interaction.	Categorical	[Yes, No]
Curb Time	Time spent at the curb: NWT (Not waiting at curb), SWT $(\leq 10s), LWT (\geq 10s)$.	Categorical	[NWT, SWT, LWT]
Exposure Time	Extra exposure time apart from crossing time: NET (No extra time), SET $(\leq 10s), LET (\geq 10s)$.	Categorical	[NET, SET, LET]
Conflict Location	The lane where the conflict occurs.	Categorical	[Shantinagar, Aloknagar]
Stage of Crossing	Indicates if the pedestrian crosses in one go or multiple stages.	Categorical	[Single, Multiple]
Conflict Type	Type of conflict: Maneuver Conflict (path change), Speed-Based Conflict (speed adjustment), Hybrid Conflict (both path and speed adjustments).	Categorical	[Maneuver Conflict, Speed-Based Conflict, Hybrid Conflict]

Table 2: Categorical Variables and Their Composition

Variable	Categories	Composition (%)
Gender	[Male, Female]	[50.4%, 49.6%]
Age	[<20, 20-50, >50]	[23.03%, 64.72%, 12.25%]
Size	[Size_1, Size_2, Size>2]	[58.71%, 25.51%, 15.76%]
Own Path	[Assertive, Reserved]	[61.41%, 38.59%]
Crossing Pattern	[Jaywalker, Normal]	[18.80%, 81.95%]
Stage of Crossing	[Single, Multiple]	[72.61%, 27.39%]
Curb Time	[NWT, SWT, LWT]	[65.97%, 17.84%, 16.18%]
Exposure Time	[NET, SET, LET]	[48.34%, 32.57%, 19.08%]
Conflict Location	[Shantinagar, Aloknagar]	[24.31%, 75.69%]
Vehicle Type	[2W, SUV-Car, Bus-Truck]	[68.65%, 28.28%, 3.05%]
Yield	[Yes, No]	[17.73%, 82.27%]
Conflict Type	[Maneuver Conflict, Speed-Based Conflict, Hybrid Conflict]	[39.44%, 37.00%, 23.54%]

Table 3: Summary of Continuous Variables

Continuous Variable	Statistics
Post Encroachment Time (PET)	[Mean = 1.19] [Std = 0.369] [Min = 0.50] [Max = 2.24]
Traffic Speed	[Mean = 9.22] [Std = 2.91] [Min = 2.54] [Max = 17.94]
Crossing Speed	[Mean = 1.19] [Std = 0.369] [Min = 0.50] [Max = 2.24]
Perceived Volume	[Mean = 67.18] [Std = 27.58] [Min = 10] [Max = 142.52]

The Box Plot of Post Encroachment Time (PET) across key variables—Vehicle Type, Pedestrian Age, and Conflict Type—revealed significant insights into pedestrian-vehicle interactions. The analysis showed that two-wheelers (2W) are associated with the lowest median PET 1.07 seconds, indicating higher risk due to their agility and rapid speed adjustments. In contrast, SUVs and cars exhibited a moderate median PET 1.30 seconds, while buses and trucks recorded the highest median PET 1.60 seconds, reflecting safer interactions attributable to their predictable movement and restricted daytime access on the ring road.

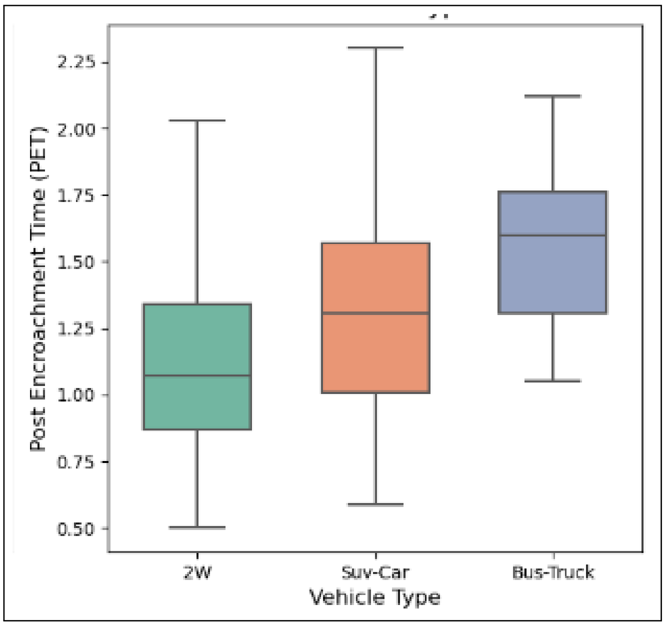


Figure 2: PET with respect to Vehicle type

When examining pedestrian age, younger pedestrians (under 20 years) demonstrated the lowest median PET (1.05 seconds), likely due to impulsive crossing behavior, making them more vulnerable to conflicts. Pedestrians aged 20–50 years showed a moderate median PET 1.15 seconds, while older pedestrians (over 50 years) exhibited the highest median PET 1.30 seconds, aligning with their cautious and deliberate crossing strategies.

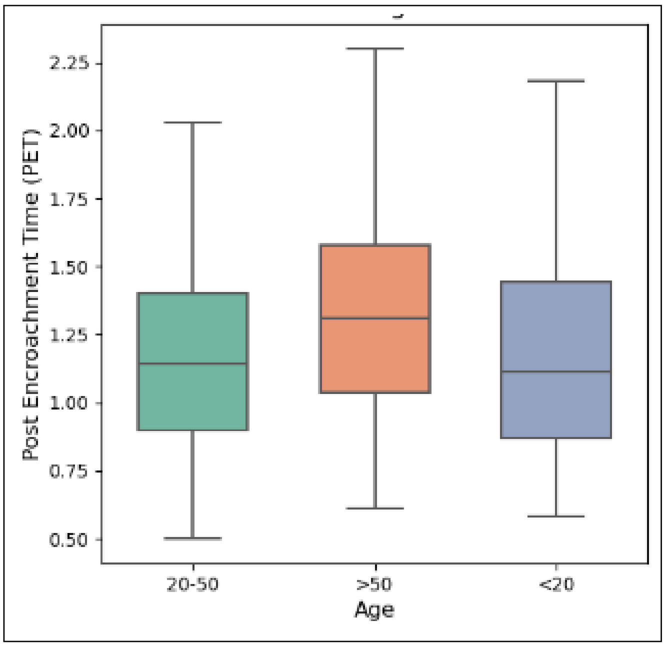


Figure 3: PET with respect to Age group

The analysis of conflict types further highlighted that hybrid conflicts—where both speed and path adjustments occur simultaneously had the lowest median PET 1.00 s, indicating the highest risk due to their unpredictable nature. Speed-based conflicts (involving speed adjustments only) showed a moderate median PET 1.16 seconds, while maneuver conflicts (involving path adjustments only) recorded the highest median PET (approximately 1.27 seconds, suggesting greater predictability and safety.

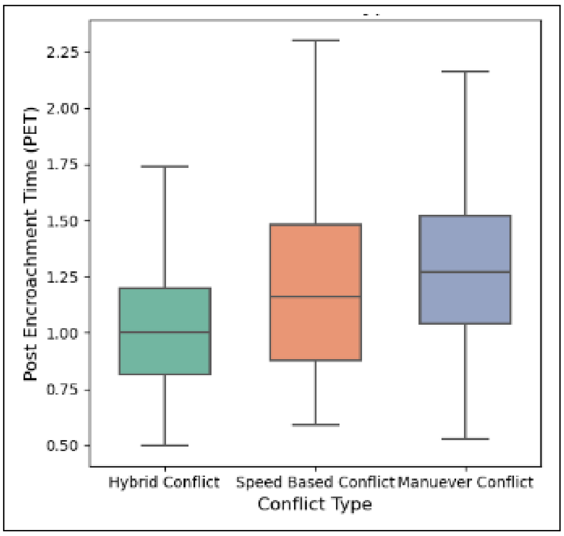


Figure 4: PET with respect to Conflict type

4.2 Correlation among variables

Pearson correlation analysis was conducted for continuous variables, while Cramér's V test was used to assess the association between categorical variables. The Pearson correlation results indicate that all correlations among independent variables are below 0.5, suggesting no significant multicollinearity. The results are visually represented in the heatmap as shown in Fig 5.

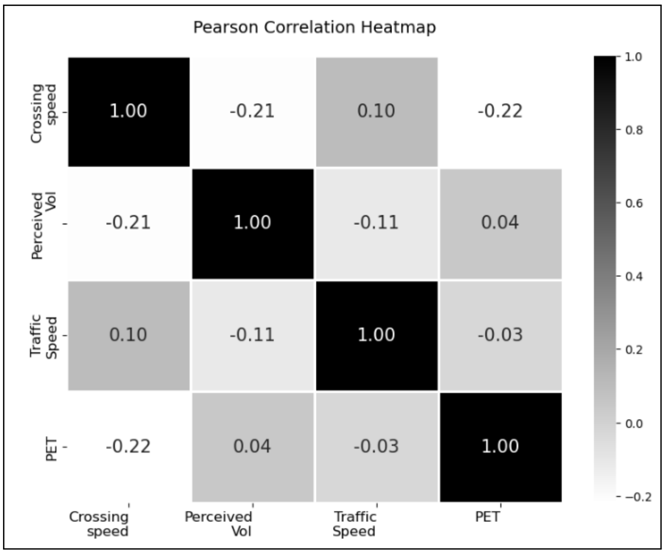


Figure 5: Corelation of Continuous variables

Similarly, in the Cramér's V test as shown in Fig 6, a notable correlation (0.61) was observed between Stage of Crossing and Exposure Time, indicating a moderate association. However, multicollinearity is typically assessed using Variance Inflation Factor (VIF) in regression modeling. Since all VIF values were below 5, it was determined that these variables did not cause significant multicollinearity issues in the model. Therefore, both were retained in the initial regression analysis to evaluate their individual contributions to PET.

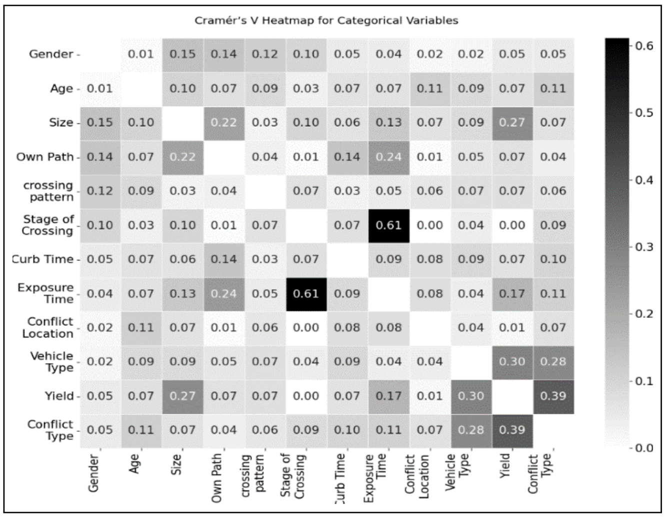


Figure 6: Corelation of Categorical variables

4.3 Multiple Linear Regression Model for Post Encroachment Time

An initial multiple linear regression model developed considering all variables explained 24.6% of the variance in PET as shown in Table 4. The p-values for each variable are reported in Table 5, and variables with p-values less than 0.05 were considered statistically significant.

Table 4: OLS Regression Results of Model including all Variables

R^2	Adj. R^2	F-Stat	Prob(F)	Log-L	AIC	BIC
0.246	0.207	6.271	9.89×10^{-13}	−103.52	243	312.2

Table 5: Feature Significance Results

Feature	P>|t|
ordinal__Age	0.492
ordinal__Size	**0.011**
ordinal__Curb Time	0.448
ordinal__Exposure Time	0.435
nominal__Gender_1	**0.021**
nominal__Own Path_1	0.348
nominal__Crossing Pattern_1	0.478
nominal__Vehicle Type_Bus-Truck	**0.000**
nominal__Vehicle Type_SUV-Car	**0.000**
nominal__Yield_1	0.095
nominal__Stage of Crossing_1	0.456
nominal__Conflict Location_1	**0.019**
nominal__Conflict Type_Manuever Conflict	**0.000**

To improve the model's interpretability and predictive power, a refined model was developed by retaining only the statistically significant variables from the initial model. The result of the model are presented in Table 6 and 7. The refined model explained 22.5% of the variance in PET ($R^2 = 0.225$), with an adjusted R^2 of 0.21, which is considered satisfactory for behavioral studies. The model's overall significance was confirmed by an F-statistic of 15.35 as shown in Table 6. Previous research has also reported relatively low R^2 values, such as in [18] (Pseudo $R^2 = 0.180$), [19] (Pseudo $R^2 = 0.503$) ,and [20] (Pseudo $R^2 = 0.306$), highlighting the inherent complexity in behavioral interactions. Given the stochastic nature of pedestrian behavior and vehicle interactions, achieving a higher R^2 in such studies remains challenging [21].

Table 6: OLS Regression Results of Model including Significant Variables

R^2	Adj. R^2	F-Stat	Prob(F)	Log-L	AIC	BIC
0.225	0.21	15.35	7.49×10^{-20}	−139.43	296.9	333.5

Table 7: Feature Coefficients and Statistical Results

Feature	Coef	Std Err	t	P|t|	0.025	0.975
const	1.12	0.03	32.66	0.00	1.05	1.19
ordinal__Size	-0.04	0.02	-1.66	0.10	-0.08	0.01
nominal__Gender_Male	-0.05	0.03	-1.50	0.14	-0.11	0.02
nominal__Vehicle Type_Bus-Truck	0.44	0.10	4.53	0.00	0.25	0.63
nominal__Vehicle Type_SUV-Car	0.22	0.04	5.78	0.00	0.15	0.29
nominal__Conflict Location_Shantinagar	-0.17	0.05	-3.56	0.00	-0.26	-0.07
nominal__Conflict Type_Manuever Conflict	0.26	0.04	6.75	0.00	0.19	0.34
numeric__Traffic Speed	-0.07	0.02	-2.97	0.00	-0.11	-0.02
numeric__Crossing speed	-0.08	0.02	-4.97	0.00	-0.12	-0.05

In the analysis of pedestrian–vehicle interactions, vehicle type emerged as a significant predictor of post-encroachment time (PET). Encounters involving Bus-Truck ($\beta = 0.439$, $p < 0.001$) and SUV-Car ($\beta = 0.219$, $p < 0.001$) were associated with higher PET, likely reflecting the larger size and slower maneuverability of these vehicles, which prolong the duration of potential conflicts. Additionally, conflict dynamics played a critical role; while maneuver conflicts were linked to increased PET ($\beta = 0.263$, $p < 0.001$), as pedestrians and vehicles adjust their paths to avoid collisions, resulting in prolonged interaction time. Speed variables also significantly affected PET—both higher crossing speed ($\beta = -0.084$, $p < 0.001$) and higher traffic speed ($\beta = -0.066$, $p = 0.003$) contributed to shorter PET, suggesting that faster movements reduce the window for conflicts. Marginal effects were observed for pedestrian group size ($\beta = -0.037$, $p = 0.098$) and male gender ($\beta = -0.049$, $p = 0.135$), though these associations were weak and not statistically significant, indicating a less robust relationship with PET.

To assess multicollinearity, Variance Inflation Factor (VIF) was calculated. A VIF above 5 indicates high collinearity, but all variables in Table 8 had VIF values below 2, suggesting no severe multicollinearity. The highest VIF (1.870 for Traffic Speed) confirms the model's stability and reliable predictor independence.

Table 8: Variance Inflation Factor (VIF) Results

Feature	VIF
ordinal__Size	1.06
nominal__Gender_Male	1.03
nominal__Vehicle Type_Bus,Truck	1.04
nominal__Vehicle Type_SUV/Car	1.12
nominal__Conflict Location_Shantinagar	1.39
nominal__Conflict Type_Manuever Conflict	1.37
numeric__Traffic Speed	1.87
numeric__Crossing speed	1.07

The Multiple Linear Regression (MLR) model was validated through statistical tests, confirming key assumptions. The Shapiro-Wilk test indicates approximately normal residuals (p = 0.14), the Residuals vs. Fitted values plot as shown in Fig 7 shows no heteroscedasticity, and the Durbin-Watson statistic (2.02, p = 0.874) confirms no significant autocorrelation. These results ensure the model's reliability for prediction.

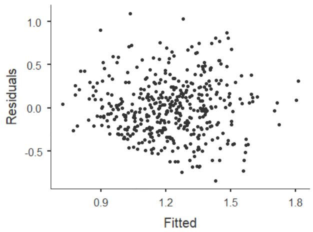


Figure 7: Residual vs Fitted Values

5. Conclusions

The analysis identifies key factors influencing pedestrian-vehicle interactions, with vehicle type and conflict dynamics playing significant roles in pedestrian safety. Larger vehicles, such as buses and trucks, are associated with longer post-encroachment times (PET), suggesting slower, more predictable interactions with pedestrians. In contrast, conflicts associated with the two-wheelers, exhibit the shortest PET, highlighting a higher risk due to their agility and rapid maneuverability. Younger pedestrians (<20 years) are at the highest risk, driven by impulsive crossing behavior. Assertive pedestrian behavior and multi-stage crossings further elevate exposure, particularly in high-conflict lane like Aloknagar side. The analysis also found that hybrid conflicts, where both speed and path adjustments occur simultaneously, are the most hazardous, accounting for approximately 23.54% of the total conflicts. The low yielding rate (17.73% of vehicles) exacerbate safety challenges. In social science research, a low R^2 value does not imply a poor model. The focus is often on assessing the significance of predictors rather than precise prediction. Even an R^2 as low as 0.1 can be acceptable provided that the predictors or explanatory variables are statistically significant [21]. This approach is particularly relevant in the context of modeling Post Encroachment Time (PET), where the focus is on understanding the relationships between variables rather than achieving high predictive accuracy. Despite the model's moderate explanatory power ($R^2 = 0.225$), the findings underscore the need for targeted interventions, such as pedestrian-activated signals, dynamic speed feedbacks, and age-specific safety programs. Stricter yielding laws, infrastructure upgrades, and awareness campaigns for drivers and pedestrians are essential to mitigate risks and enhance safety in high-conflict areas.

6. Recommendations

To enhance the predictive accuracy of the model, several additional variables and advanced modeling techniques could be considered. Incorporating the number of interacted vehicles within a crosswalk would provide insight into the cumulative risk faced by pedestrians, particularly in high-traffic conditions. Including the accepted gaps—defined as the time pedestrians are willing to wait between vehicles—would capture crossing behavior and better reflect risk exposure. Considering pedestrian familiarity with the crossing location could provide insights into behavioral differences between regular and unfamiliar pedestrians. Beyond expanding the dataset with these variables, exploring other machine learning and non-linear models could significantly improve the model's performance. Non-linear models like decision trees, random forests, or gradient boosting machines can capture complex, non-linear relationships between variables that traditional linear regression may miss.

References

[1] WHO. Global status report on road safety 2023. 2023.

[2] Rajendra Karkee and Andy H Lee. Epidemiology of road traffic injuries in nepal, 2001–2013: systematic review and secondary data analysis. *BMJ Open*, 6:e010757, 4 2016.

[3] Salaheddine Bendak, Asayel M. Alnaqbi, Muna Y. Alzarooni, Sara M. Aljanaahi, and Shaikha J. Alsuwaidi. Factors affecting pedestrian behaviors at signalized crosswalks: An empirical study. *Journal of Safety Research*, 76:269–275, 2 2021.

[4] P. P. Koh, Y. D. Wong, and P. Chandrasekar. Safety evaluation of pedestrian behaviour and violations at signalised pedestrian crossings. *Safety Science*, 70:143–152, 12 2014.

[5] Leah L Thompson, Frederick P Rivara, Rajiv C Ayyagari, and Beth E Ebel. Impact of social and technological distraction on pedestrian crossing behaviour: an observational study. *Injury Prevention*, 19:232–237, 8 2013.

[6] András Várhelyi. Drivers' speed behaviour at a zebra crossing: a case study. *Accident Analysis Prevention*, 30:731–743, 11 1998.

[7] Sumona Bajracharya and Anil Marsani. Study on drivers' yielding behavior at pedestrian crossings. pages 2350–8906, 2021.

[8] RAY FULLER. A conceptualization of driving behaviour as threat avoidance. *Ergonomics*, 27:1139–1155, 11 1984.

[9] Matus Sucha, Daniel Dostal, and Ralf Risser. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis Prevention*, 102:41–50, 5 2017.

[10] Ralf Risser. Behavior in traffic conflict situations. *Accident Analysis Prevention*, 17:179–197, 4 1985.

[11] Matúš Šucha. Road users' strategies and communication: driver-pedestrian interaction, 2014.

[12] Carl Johnsson, Aliaksei Laureshyn, and Tim De Ceunynck. In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport Reviews*, 38:765–785, 1 2018.

[13] Ashutosh Arun, Md Mazharul Haque, Simon Washington, Tarek Sayed, and Fred Mannering. A systematic review of traffic conflict-based safety measures with a focus on application context. *Analytic Methods in Accident Research*, 32:100185, 12 2021.

[14] Lai Zheng, Tarek Sayed, and Mohamed Essa. Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. *Accident Analysis Prevention*, 123:314–323, 2 2019.

[15] Ahmed Tageldin, Tarek Sayed, and Khaled Shaaban. Comparison of time-proximity and evasive action conflict measures: Case studies from five cities. *Transportation Research Record: Journal of the Transportation Research Board*, 2661:19–29, 1 2017.

[16] Mohamed Gomaa Mohamed and Nicolas Saunier. The impact of motion prediction methods on surrogate safety analysis: A case study of left-turn and opposite-direction interactions at a signalized intersection in montreal. *Journal of Transportation Safety Security*, 10:265–287, 7 2018.

[17] Lakshmi N. Peesapati, Michael P. Hunter, and Michael O. Rodgers. Can post encroachment time substitute intersection characteristics in crash prediction models? *Journal of Safety Research*, 66:205–211, 9 2018.

[18] P. Vedagiri and B. Raghuram Kadali. Evaluation of pedestrian–vehicle conflict severity at unprotected midblock crosswalks in india. *https://doi.org/10.3141/2581-06*, 2581:48–56, 1 2016.

[19] Deepak Raj Shah and Rojee Pradhananga. Proceedings of 12 th ioe graduate conference a study on the factors influencing red light violation behavior of pedestrians in kathmandu valley. page 12.

[20] B. Raghuram Kadali and P. Vedagiri. Proactive pedestrian safety evaluation at unprotected mid-block crosswalk locations under mixed traffic conditions. *Safety Science*, 89:94–105, 11 2016.

[21] Peterson K Ozili. The acceptable r-square in empirical modelling for social science research. *SSRN Electronic Journal*, 6 2022.