UHASSELT
KNOWLEDGE IN ACTION

School of Transportation Sciences

Master of Transportation Sciences

Master's thesis

EFFECTIVENESS OF I-DREAMS TECHNOLOGY IN HIGH vs LOW-RISK ZONES

Getahun Asres Belay
Thesis presented in fulfillment of the requirements for the degree of Master of Transportation Sciences, specialization Traffic Safety

SUPERVISOR :
Prof. dr. Kris BRIJS

MENTOR :
De heer Muhammad Wisal KHATTAK

2023
2024

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

School of Transportation Sciences

Master of Transportation Sciences

Master's thesis

EFFECTIVENESS OF I-DREAMS TECHNOLOGY IN HIGH vs LOW-RISK ZONES

Getahun Asres Belay

Thesis presented in fulfillment of the requirements for the degree of Master of Transportation Sciences, specialization Traffic Safety

SUPERVISOR :

Prof. dr. Kris BRIJS

MENTOR :

De heer Muhammad Wisal KHATTAK

ABSTRACT

Road safety remains a significant public health concern. Despite many advancements in vehicle technologies, there is still a pressing need for research initiatives like i-DREAMS, which aim to reduce road fatalities and injuries by analyzing risky driving events. Risky driving events (Speeding, Acceleration, Deceleration, Tailgating, Steering, etc.) are near-crash incidents where a crash is narrowly avoided and provide critical insights into driving behaviors that may not be captured through crash data alone. Traditionally, road safety studies have relied on crash data to assess safety levels and evaluate intervention effectiveness. However, recent research has shifted towards using naturalistic driving data to capture and analyze critical incidents, providing valuable insights into driver behavior during near-crash events.

The i-DREAMS intervention is designed to enhance road safety by monitoring and improving driver behavior. It uses advanced technology to assess the driving environment, vehicle dynamics, and driver status, delivering timely interventions to prevent risky driving behaviors (Brijs et al., 2020). The i-DREAMS platform aims to keep drivers within a "Safety Tolerance Zone" (STZ) by analyzing risky driving incidents and providing real-time in-vehicle warning and post-trip feedback, thereby reducing the likelihood of accidents. This study employed spatial and statistical analysis to assess the effectiveness of the i-DREAMS intervention in reducing risky driving behaviors, specifically tailgating and speeding, across different risk zones and roadway categories.

The spatial analysis identified high-risk zones and hotspots for these events, primarily concentrated in major cities and high-traffic areas in the Limburg province of Belgium. Notable hotspots were centered around Hasselt and Genk, predominantly on motorways and primary roads.

The Generalized Estimating Equations (GEE) with a Negative Binomial Model and Log link function were used to analyze the data. The intervention phases generally reduced the number of tailgating and speeding events from the baseline phase. However, the statistical analysis revealed that only some reductions reached statistical significance. Additionally, no significant differences were found in the rate of change between high- and medium-severity events for both behaviors. The statistical comparison of the effectiveness of the intervention between high and low-risk zones revealed that the intervention was significantly more effective in high-risk zones for tailgating events, with no significant impact in low-risk zones. Moreover, neither high-risk nor low-risk zones showed significant improvements in speeding events. Further, the effectiveness did not vary significantly across different road types.

These findings illustrate that the i-DREAMS intervention effectively reduced tailgating events, particularly in high-risk zones, indicating that the intervention phases had a significant impact on mitigating this specific risky driving behavior in areas where it was most prevalent. The insignificance of the intervention's overall impact on speeding events also suggested that the strategies employed were less effective in addressing speeding behavior. Moreover, this study contributed to identifying which risky driving behaviors and risk zones the i-DREAMS intervention is most effective for and highlighting where the intervention strategies may need redesigning to address specific challenges and improve overall effectiveness.

Keywords: Speeding, Tailgating, i-DREAMS, Spatial Analysis, Negative binomial Model, Naturalistic Driving Study

ACKNOWLEDGMENT

First and foremost, I express my deepest gratitude to God for His endless mercy, granting me excellence and strength during times of difficulty and weakness. I sincerely thank my research supervisor, Prof. Dr. Kris Brijs, for his genuine guidance and invaluable comments in achieving my research goals. I am also profoundly grateful to my mentor, Mr. Wisal Khattak, for his insightful feedback and unwavering guidance throughout this research. I sincerely appreciate the i-DREAMS project research team for providing me with the necessary data necessary for this study.

I am also deeply thankful to my colleagues and fellow students at Hasselt University for their solidarity and support. I am also immensely grateful to my family for their unwavering encouragement and motivation. Last but not least, I would like to acknowledge the Flemish Inter-University Council (Vlaamse Interuniversitaire Raad/VLIR-UOS) for awarding me a scholarship, which enabled me to pursue my master’s studies at Hasselt University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 General Background

Road traffic safety is one of the most significant challenges that the world deals with, and the United Nations (UN) has dedicated considerable resources to address and mitigate this public health issue. According to the Global Status Report on Road Safety 2023, annual road traffic deaths have slightly decreased to 1.19 million. However, Road traffic injuries remain the leading killer of children and young people aged 5-29 years. More than half of fatalities occur among pedestrians, cyclists, and motorcyclists, in particular those living in low and middle-income countries. The United Nations (UN) develops action plans to lower road fatalities every ten years. Despite some progress, the economic impact of road crashes remains significant, with losses estimated at about 3% of GDP in affected countries. The report highlights the need for sustained efforts and proven measures to achieve the global goal of halving road traffic deaths and injuries by 2030 (WHO, 2023).

In 27 EU countries, there were 18,844 traffic-related fatalities in 2020, 37% fewer compared to 2010. The goal of halving crashes by 2020 was not met, and the declining rate has usually slowed since 2013 (Adminaité-Fodor et al., 2021). This shows that the implemented countermeasures and intervention strategies were less effective in reducing vehicle occupant fatalities and injuries. Accordingly, cutting-edge and innovative research efforts and initiatives are crucial to achieving the ambitious long-term objective of decreasing road fatalities and severe injury collisions in the EU to zero by 2050 (European Commission, 2020). To overcome this problem, the EU set interim-term and long-term goals. The interim target of the EU is to reduce road fatalities and severe injury by 50% from 2020 to 2030. Similarly, the Flemish area of Belgium aims for zero deaths by 2050 after cutting them in half by 2030 (DANIELS, 2022). Achieving these goals requires the coordinated effort of all transport stakeholders, including government agencies, road safety organizations, vehicle manufacturers, infrastructure planners, and the general public. Thus, each player must play their part in a coordinated way to accomplish these objectives and efforts. Road safety research has been consistently highlighting multiple factors contributing to road traffic crashes. This issue was first addressed by the Haddon Matrix (Haddon, 1968), an epidemiological approach to traffic safety by categorizing the factors into three phases (pre-crash, crash, and post-crash) and across three domains: human, vehicle, and environment. This approach enables a systematic identification of potential countermeasures across each phase and domain. For example, in the pre-crash phase, human factors might include driver education and enforcement of traffic laws, while vehicle factors could involve the implementation of advanced safety technologies like anti-lock braking systems. Environmental factors in this phase might focus on road design improvements. Generally, human Factors are one of the core factors in road crashes. Even if these human factors contribute to more than 90% of road crashes, researchers have given them little attention (Shaon, 2019). Among all human factors, driver behavior is the key to safe mobility. Most road crashes are caused by risky driving behavior, characterized by risk events such as speeding, decelerating, lane changing, and tailgating (Katrakazas et al., 2020b).

Several efforts have been made to minimize the impact of human error in road accidents by introducing both partial and full automation technologies, such as Advanced Driver Assistance Systems (ADAS). These systems are designed to enhance vehicle safety by offering features like anti-lock braking systems, lane departure warnings, lane-keeping assistance, and forward collision warnings (Jayan et al., 2021). For instance, a study by Leslie (2021) revealed that Automatic Emergency Braking (AEB), camera-based Automatic Emergency Braking, and Forward Collision Alert systems were effective in reducing rear-end collisions, with reductions of 45%, 38%, and 20%, respectively. Additionally, the study found that Lane Keep Assist with Lane Departure Warning and Lane Departure Warning systems on their own contributed to decreases in lane departure crashes by 12% and 10%, respectively. Research on the acceptance of retrofitted ADAS packages, which include systems like forward collision warning and lane-keep assist, indicated that users generally have a positive view of these technologies and they appreciated these systems' perceived benefits, ease of use, and reliability (Khattak et al., 2024).

Despite these advancements in ADAS, research shows that Advanced Driver Assistance Systems (ADAS) do not effectively address long-term, stable behavioral factors like personality, driving experience, and safety attitudes, which are crucial for determining how drivers respond in real-world situations (Brijs et al., 2023). To bridge this gap, the i-DREAMS project, funded by the European Horizon2020 initiative, developed a framework for a context-aware 'safety tolerance zone (STZ)' for driving. This framework aimed to incorporate these stable behavioral factors to enhance overall driver safety (Brijs et al., 2020). Moreover, it is designed to address the road safety challenges in the digitization and automation of transportation and contributes to enhanced safety (i-DREAMS, 2019).

This study is based on the data from the i-DREAMS project to examine the impact of the i-DREAMS intervention on risky driving behaviors, focusing on specific events such as speeding and tailgating. The study considered various contexts, including the severity of the events, risk zones, and different road categories. The i-DREAMS project, a naturalistic driving study, has been conducted across five European countries, including Belgium (Talbot et al., 2020).

1.1.1 Background on the i-DREAMS Project

The i-DREAMS project is funded by the EU and coordinated by the transportation research institute (IMOB), Hasselt University, to define, develop, test, and validate a ‘Safety Tolerance Zone (STZ).’ The STZ is a context-aware safety envelope that prevents drivers from getting too close to unsafe operations by mitigating risks through in-vehicle and post-trip interventions. The STZ comprises three distinct driving phases: ‘normal,’ ‘danger,’ and ‘avoidable crash’ (Katrakazas et al., 2020a); the normal driving phase signifies conditions where a crash is unlikely, indicating a low crash risk. In this phase, the driver effectively adjusts their behavior to match the task demands, maintaining a balance between their coping capacity and the complexity of the task. The danger phase is marked by deviations from normal driving that suggest an increased crash risk. Lastly, the avoidable crash phase arises when a collision scenario emerges, yet there is still an opportunity for the driver to take action and prevent the crash.

The main objective of the i-DREAMS platform is to keep the driver in the normal driving phase for as long as possible. If maintaining this phase is impossible, the platform system integrates real-time and post-trip interventions to prevent the progression from the danger phase to the avoidable accident phase through nudging and coaching the driver. The urgency for intervention is greater than in the danger phase, and if the driver fails to modify their behavior accordingly, a crash is highly probable.

The i-DREAMS monitors the safety level of drivers’ process and intervenes in real-time and post-trip, considering ‘operator,’ ‘vehicle,’ and ‘context-related’ measurements (Talbot et al., 2020). In real-time, based on the imminent crash risks, in-vehicle interventions such as a warning signal are issued to assist and support vehicle drivers. Post-trip interventions are provided after the trip is over, and they are specifically molded to each operator’s driving style and the number of risky events encountered. This intervention is further classified into two sub-phases. The first is the provision of feedback on a smartphone, and the second includes post-trip feedback on a smartphone and a gamified web platform designed to improve their driving gradually (Papadimitriou et al., 2019; Katrakazas et al., 2020).

The framework of the i-DREAMS project is being tested using a simulator study and three stages of on-road trials in five countries in Europe: Germany, Greece, Portugal, Belgium, and the UK. The test trial involves 600 drivers distributed to the five countries and four types of transport modes: cars, trucks, buses, and rail. The simulators are used to test driving behaviors, the i-DREAMS equipment like Cardio Wheel and Mobileye, and the technologies developed for real-time interventions. The on-road trials are classified into pilot testing, baseline (no intervention), and intervention stages. The intervention stage is organized into three levels: in-vehicle intervention, post-trip feedback on smartphones, and post-trip feedback on smartphones and gamified web platforms (Hancox, 2020).

1.1.2 Intervention Mapping in the i-DREAMS Project

Intervention mapping (IM) is a methodical approach to creating behavioral interventions, and it offers a step-by-step guide to developing a theory and evidence-based intervention program (Gitlin & Czaja, 2015). The six steps of intervention mapping include the logic model of the problem, logic model of change, program design, program production, program implementation plan, and evaluation plan (Gitlin & Czaja, 2015). The i-DREAMS project has adopted the principles of IM to guide its decision-making process throughout the development and operationalization of its interventions. Using the IM, a toolbox that guides planners in developing real-time and post-trip interventions is prepared, and it covers the aspects of the selection of devices, how the user interface of the app and the web platform look like, and what kind of information and how it should be provided, etc. (Brijs, 2020).

The first step of IM involves constructing a logic model that identifies the road safety issues among car and professional drivers involved in accidents, pinpointing specific safety outcomes, goals promoting safety, risky behaviors, and their determinants (Brijs, 2020). This model outlines five safety outcomes expected from real-time and post-trip interventions: substantial reductions in frontal, side, rear, roll-over/derailment crashes, and passenger injuries in cars, buses, trucks, and trams. The associated safety-promoting goals are logically derived from the behaviors linked to these outcomes, such as improved vehicle control, road-sharing, speed management, and driving when adequately fit. These goals are then broken down into more specific risky behaviors, which are transformed into performance objectives, detailed in Table 1.1. This logical analysis culminates in pinpointing the determinants addressed by the interventions (Brijs, 2020).

Table 1.1: SPGs, Performance objectives, and the i-DREAMS metric (Katrakazas et al., 2020c)

Safety promoting goal	**Performance objectives**	**I-DREAMS metric**
Vehicle control	Acceleration	Number of harsh accelerations and acceleration aggressiveness level
	Deceleration	Number of harsh decelerations and deceleration aggressiveness level
	Steering	Number of harsh cornering
Sharing the road with others	Tailgating	Number of headways monitor warnings
	Lane Discipline	Number of lane departure warnings

	Overtaking	Number of illegal overtaking events
	Forward collision avoidance	Number of forward-collision warnings and urban forward-collision warnings
	Lane departure avoidance	Number of lane departure warnings and dashcam video
	Vulnerable road user collision avoidance	Number of pedestrian collision warning events
Speed management	Speeding (Exceeding the speed limit)	Percentage of overseeding and average speed over the speed limit
Driving under a healthy condition	Fatigue	KSS-score
		Driving during the night and trip duration
	Distraction	Handheld mobile phone use
		No hands on the wheel
Using safety devices	----	Number of tamper alerts

The second step involves constructing a change logic model by creating a matrix that illustrates the relationships between change objectives, performance objectives, and safety-promoting goals. Each safety-promoting goal (SPG) generates specific change objectives (COs) based on its performance objectives and their determinants, such as enhancing drivers' road-sharing by reducing risky tailgating, with attention as a key determinant for real-time interventions (Brijs, 2020). Table 1.1 presents these elements, including the i-DREAMS metrics used to measure them (Katrakazas et al., 2020c). Each SPG directly corresponds to crucial driving behaviors linked to desired safety outcomes, with performance objectives operationalizing the change required and i-DREAMS metrics quantifying them.

The third step in the i-DREAMS IM strategy involves designing intervention methods that achieve the COs established in the second step. This involves theoretical and practical analyses to develop a model of change, incorporating both real-time and post-trip interventions. In the real-time setting, interventions are activated based on the safety levels determined by the Safety Tolerance Zone (STZ), with different signals used for normal driving zones compared to danger and avoidable accident phases, often through nudging techniques. Post-trip interventions focus on influencing safety-related attitudes and driving skills, adopting a coaching style (Brijs, 2020).

The fourth step, intervention production, involves creating prototypes or mock-ups for these interventions. Different symbols and sounds are designed for real-time interventions for each STZ phase, whereas post-trip interventions utilize a smartphone app or a combination of smartphone and web platforms (Brijs, 2020). Finally, the fifth and sixth steps of the IM program implementation and evaluation follow, ensuring the successful delivery and assessment of the i-DREAMS intervention program. The evaluation encompasses various parameters, including the interventions' efficacy, implementation and uptake, and economic impact (Katrakazas et al., 2020c). These steps of the intervention mapping aim to improve the outcomes proposed in the logic model of change presented in Figure 1.1.

Figure 1.1: The logic model of change in the i-DREAMS intervention (Brown et al., 2023)

1.1.3 On-road Trials Testing of the i-DREAMS project

To test its framework, the i-DREAMS project employs a simulator study and on-road trials across five European countries, including Belgium. The on-road trials are segmented into pilot testing, a baseline stage without interventions, and an intervention stage with three levels: in-vehicle interventions, post-trip smartphone feedback, and gamified web platform feedback (Hancox, 2020).

The i-DREAMS platform utilized two key modules. The first was the monitoring module, which gathered data related to the driving context, the driver, and the vehicle to assess the demands of the driving task and the driver's capacity to meet these demands. This module is used to determine the driver's current phase within the Safety Tolerance Zone (STZ) at any given moment. The second module, the intervention module, aimed to keep the driver within the normal phase of the STZ by providing either real-time warnings or alerts during the trip (real-time intervention) or feedback after the trip (post-trip intervention). The type of in-vehicle warning delivered varies based on the severity of the detected event (Brown et al., 2023).

Drivers received real-time warnings through the in-vehicle i-DREAMS display whenever they entered the 'danger phase' (STZ level 2) for a particular Performance Objective (PO). A more intrusive warning was issued if the situation escalated to the 'avoidable crash phase' (STZ level 3). These warnings included both visual and audible alerts, as detailed in the 'Real-time interventions manual' in the i-DREAMS Deliverable 5.3 (Hancox et al., 2021). The real-time warnings addressed POs related to 'speeding,' 'road sharing,' and 'driver fitness,' but there were no warnings for POs associated with 'vehicle control' (Hancox et al., 2021).

Post-trip feedback was provided to drivers through the i-DREAMS smartphone app, which offered various functions during both Phase 3 and Phase 4 of the trial, described in detail in Deliverable 4.5 (Vanrompay et al., 2020). Drivers could access 'Scores' to view trip-specific and aggregated scores for each Performance Objective (PO) and Safety Performance Goal (SPG), track changes over time, and analyze details of their trips through the 'Trips' feature, including visualizing trips on a map with event markers and accessing video recordings for some events. During Phase 4, additional features included informational 'Tips' related to safe driving, a 'Leader board' for comparing scores with other drivers, and 'Goals and badges' as part of the gamification element, where drivers could earn virtual badges for achieving and maintaining specific driving goals.

1.2 Problem Statement

Many studies have shown the effectiveness of in-vehicle and post-trip interventions in reducing unsafe driving behaviors and their consequences. In-vehicle interventions, such as providing information on the observed risky behavior in real-time during the trip, have been studied for their effectiveness in reducing excessive speeding effectively (Zhao & Wu, 2013), tailgating (Adell et al., 2011), and distracted driving (Aidman et al., 2015). However, some studies show that drivers show no improvement or worsen in other driving behaviors such as lane changing, speeding, and overtaking (Adell et al., 2011).

The effectiveness of in-vehicle and post-trip interventions can also vary depending on the location of the targeted risky events. For example, feedback on driving behavior is more effective in urban areas, where there are more opportunities for drivers to engage in risky behavior (Kircher et al., 2007). Similarly, roadside advertising signs have been found to have a more significant impact on driver behavior in rural areas, where there are fewer distractions on the road (Keall et al., 2004). Moreover, studies suggest that in-vehicle interventions can effectively reduce risky events in different risk zones.

However, it is essential to note that the effectiveness of these interventions can vary depending on the specific intervention, event severity levels, the risk zone, and other contextual factors such as road types. Moreover, there is a gap in existing literature in employing spatial analysis tools to identify high-risk and low-risk zones, providing a more localized understanding of where interventions could be most effective.

Specifically, the i-DREAMS outcome evaluation also aimed to assess the impact of real-time driver interventions, post-trip feedback, and gamification on driving behavior and driver state. This evaluation intended to focus on safety-promoting goals, performance objectives, and change objectives, acknowledging that detecting the interventions' impact on rare events like crashes is unlikely due to the short duration of the field trials (Katrakazas et al., 2020c). Based on the motivation of filling these gaps, this study is designed to evaluate and compare the effectiveness of i-DREAMS in-vehicle and post-trip interventions in lowering the frequency of risky driving events in high and low-risk zones.

This study compared the three project-developed interventions for reducing speeding and tailgating driving behaviors in high- and low-risk zones. Moreover, the study investigated the effectiveness of the interventions across different road types.

1.3 Objectives and Research Questions

1.3.1 General Objective

The general objective of this study is to compare the effectiveness of the i-DREAMS intervention in high- and low-risk zones, considering the rate at which risky events evolve from the naturalistic driving situation through the intervention's various phases using secondary data from the i-DREAMS project.

1.3.2 Specific Objectives

The specific objectives are the following.

- Identify the high- and low-risk zones in the study area, filter the associated driving risk events in those zones, and further filter the data along the different road categories, including motorways, primary roads, and secondary roads.
- Evaluate the impact of the three phases of the intervention (in-vehicle, post-trip, and post-trip smartphone and web-based gamified) in reducing the rate of risky driving events in both high and low-risk zones.
- Examine whether the effectiveness of the interventions in high and low-risk zones depends on the type of risk events.

- Examine the effectiveness of the interventions in different roadway categories
- Determine specific measures that can be taken to improve the intervention in the zone where it is less effective.

1.4 Hypothesis and Research Questions

1.4.1 Hypothesis

Statistical modeling was used to test the null and alternative hypotheses of the objectives of this study. The null hypothesis for this study was that the effectiveness of the i-DREAMS intervention is not dependent on the type of risk zone. On the other hand, the alternative hypothesis was that the effectiveness of the i-DREAMS intervention is different in high- and low-risk ones.

Ho: There is no significant difference in the effectiveness of i-DREAMS intervention in reducing risky events between high and low-risk zones.

H1: i-DREAMS intervention is more effective in reducing risky events in high- or low-risk zones.

1.4.2 Research Questions

The research questions that this study aims to answer are the following.

- How does the spatial distribution of risky driving events look in the study area?
- Do the in-vehicle, smartphone-based post-trip, and smartphone and web-based gamified post-trip interventions lower drivers' risk events rate compared to the baseline (no-intervention) condition in both high and low-risk zones?
- How effective are the in-vehicle, smartphone-based post-trip, and smartphone and web-based-gamified post-trip interventions on reducing the rate of driving risk events?
- In which zones (high or low-risk zones) are the different phases of the intervention (the in-vehicle, smartphone-based post-trip, and smartphone and web-based-gamified post-trip interventions) more effective in dropping the rate of driving risk events?
- Does the impact of the interventions differ in different roadway categories?
- What measures should be taken to improve the effectiveness of the interventions in the area where they are less effective?

1.5 Scope and Limitation of the Study

The scope of this study is specifically limited to on-road trial data of car vehicles in Belgium among the four transport modes and five European countries involved in the i-DREAMS project. The study evaluated selected performance objectives of the i-DREAMS project and did not assess other levels, such as safety outcomes, safety promoting goals, and change objectives of the logic model of change. The study is focused on identifying the high and low-risk zones in the study area and assessing the effectiveness of the i-DREAMS intervention in reducing tailgating and speeding driving events in high and low-risk zones for car vehicles in Belgium. The analysis specifically compared the intervention's impact in high and low-risk zones and across different road categories, including motorways, primary roads, and secondary roads. One significant limitation of this study is the assumption that traffic volume remained constant across all phases. This assumption was necessary due to the absence of actual traffic data for the different phases, which could affect the precision of the study's findings and limit the generalizability of the results to other settings where traffic patterns differ. Moreover, this study did not include questionary analysis and driving simulator studies of the i-DREAMS project. Furthermore, the study did not consider the effects of weather, time, or socio-demographic characteristics.

1.6 Further Research that can be made

The researcher encourages others to work on the stated limitation with a broader study area and type of risky driving events considering the weather, time, and other socio-demographic factors.

1.7 Significance of the Study

The study plays a vital role in proactive safety measures by examining the effectiveness of interventions in different risk zones within the study area through informing on areas and risky driving events where the intervention needs more refinement. It is essential for lowering risky events in various risk zones and has the potential to save lives, reduce injuries, and improve the overall safety and efficiency of roadways. In addition, the effectiveness of in-vehicle and post-trip interventions in different risk zones is studied to help inform the development of targeted interventions tailored to specific driving events and situations. This can further improve the effectiveness and efficiency of road safety interventions and help prioritize resources where they are most needed.

Moreover, this study has the potential to inform the i-DREAMS interventions by highlighting which interventions work best in high-low-risk areas and suggesting ways to improve the effectiveness of interventions when they are less successful in a particular risk zone and risky driving events. Generally, the significance of this study extends to reducing risky driving events and improving safety for all road users.

1.8 Organization of the Study

The study on the effectiveness of i-DREAMS interventions in high vs. low-risk zones is structured as presented in the chart in Figure 1.2. It begins with an Introduction that sets the stage for the research, followed by a Literature Review encompassing both local and international studies to establish a solid foundation. The Research Methodology and Materials section details the methods and materials used. This is followed by an in-depth Data Analysis that evaluates the interventions' impact. The Result and Discussion section then presents the findings in detail. Finally, the study presents a Conclusion and Recommendation, summarizing key insights and offering actionable recommendations based on the analysis.

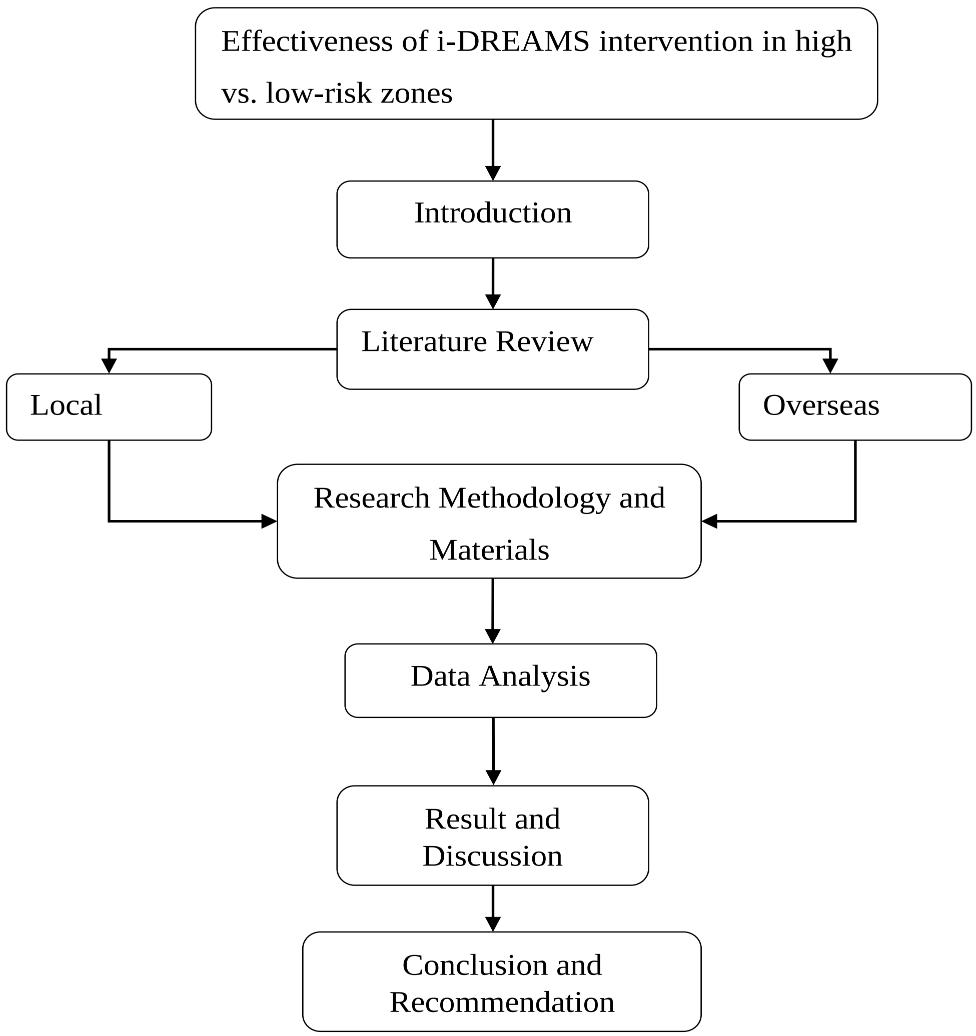

Figure 1.2: Study Structure

2. LITERATURE REVIEW

2.1 Introduction

This chapter provides an overview of the current traffic safety literature on the topic and discusses key concepts, including naturalistic driving studies, the relationship between risky driving behavior and traffic crashes, and spatial and statistical analysis methods. The chapter aims to review several relevant studies, with the following structure: Section 2.2 discusses road safety and its impacts, Section 2.3 covers naturalistic driving data analysis, Section 2.4 examines road infrastructure-related variables and risky driving events, and Section 2.5 addresses driver-related variables and risky driving events. Section 2.6 explores the relationship between risky driving behavior and traffic crashes, Section 2.7 delves into spatial analysis, Section 2.8 reviews the different approaches for evaluating intervention effectiveness, Section 2.9 discusses statistical methods for evaluating and comparing interventions, and Section 2.10 summarizes the literature review.

2.2 Road Traffic Safety

Road safety remains a critical public health concern, necessitating continuous research initiatives like i-DREAMS, which aim to reduce road fatalities and injuries by analyzing near-crash events. Despite various countermeasures, road safety has become a critical concern globally because of the continuous increase in fatalities and serious injuries.

Improving the understanding of crashes can help design effective countermeasures. Road safety become a top priority issue for all actors (government agencies, road safety organizations, vehicle manufacturers, transport infrastructure designers, and researchers) for decades. Researchers proposed many solutions for road safety problems, such as intelligent transportation systems, Autonomous Vehicles, and improved infrastructure safety by considering road user behaviors (Tarko et al., 2009). Even if many mitigations have been implemented to reduce traffic safety problems, a significant number of fatalities have occurred worldwide, and the issue still requires further improvement. Traffic safety research classifies crash causation factors into three categories: road, vehicle, and human factors. The human factor is generally considered as the primary contributor to road crashes, accounting for more than 90 % of total crashes (Shaon, 2019). Thus, managing human factors efficiently and effectively reduces traffic safety problems.

Despite significant road safety advancements, developed and middle-income countries continue to face the ongoing challenge of road traffic fatalities, although there is more progress in reducing road traffic deaths in middle-income and high-income countries than in low-income countries. Europe and other middle-income countries have made more progress in reducing road traffic deaths (WHO, 2023). In the European Union (EU), approximately 70 people still lose their lives daily in road accidents, underscoring the need for continued efforts to reduce road fatalities and serious injuries significantly (European Commission, 2020). Belgium, in particular, has made noteworthy progress, with road fatalities decreasing by 38% from 2012 to 2021, a reduction significantly higher than the EU average decrease of 25%, as presented in Figure 2.1. However, road safety remains a critical concern as 516 people were killed and 3,098 were seriously injured in road crashes in Belgium in 2021 (Folla, 2023: European Commission). The country recorded a road mortality rate of 45 fatalities per million inhabitants in 2021, which matches the EU average. Although serious injuries have decreased by 35% over the same period, these figures highlight that road traffic deaths and injuries remain a significant societal issue that requires ongoing attention and intervention (Folla, 2023: European Commission).

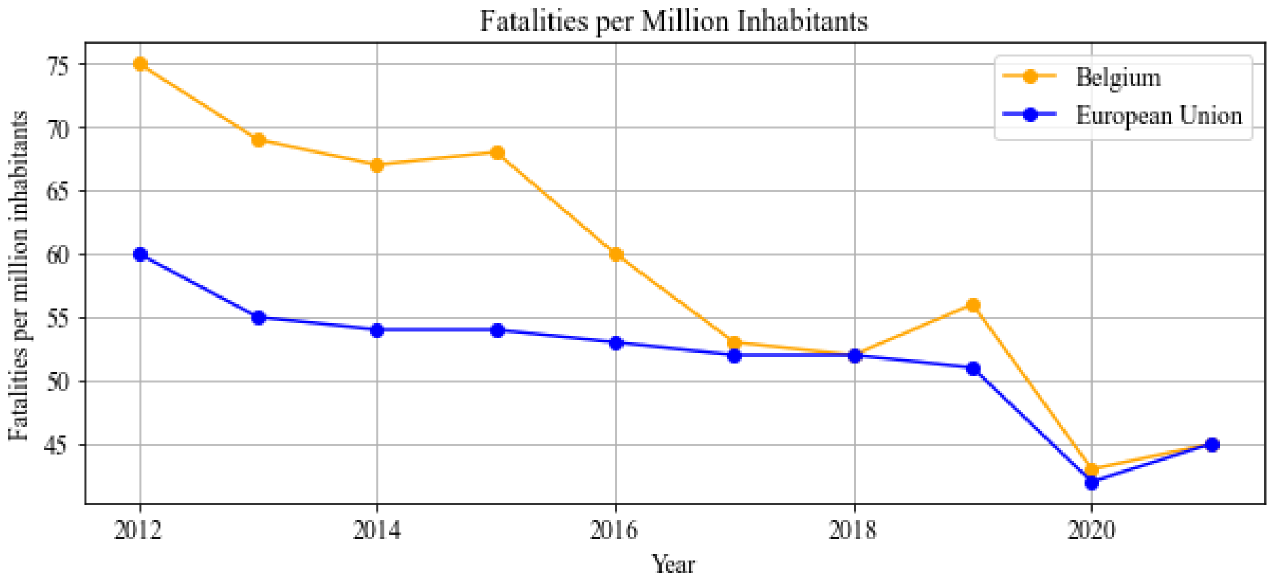

Figure 2.1: Traffic mortality rate development, 2012 – 2021 (European Commission, 2023)

2.3 Naturalistic Driving Data Analysis

Over the past decade, researchers have largely focused on analyzing actual traffic crashes in road safety studies. However, accumulating sufficient crash data for comprehensive analysis is time-consuming. Recognizing these limitations, a new approach has emerged in traffic safety research. Naturalistic Driving studies began in the early 21st Century. In naturalistic driving study, vehicles are equipped with several cameras and sensors. Sensors and cameras collect maneuvering data such as vehicle acceleration, deceleration behavior, speed management, tailgating, road sharing, and drivers' distraction or fatigue condition through sensing drivers' eyes, heads, and hands. In addition, the external conditions are recorded by naturalistic studies, such as road, traffic, and weather conditions (Schagen et al., 2012).

The naturalistic Driving study opens an environment to understand the relationship between driver's behavior, vehicle, road infrastructure, traffic, and weather conditions associated with crashes and road safety problems. In Europe, Several Naturalistic Driving Studies were undertaken, primarily focusing on road traffic, such as INTERACTION (studies of the interaction of drivers with mobile phones and various navigation systems), DaCOTA (studies on the feasibility of Naturalistic Driving for monitoring performance indicators and exposures), 2BE-SAFE (studies about motorcycle safety) and PROLOGUE (Promoting accurate life observation for gaining an understanding of road user behavior in Europe which conducted between 2009 to 2011) (Schagen et al., 2012).

Crash surrogates are non-crash events but are used as indicators to analyze and predict potential crash occurrences. These surrogates represent intermediate states between safe driving conditions and a crash, offering insights into the factors that could lead to a crash (Tarko et al., 2009). Crash surrogates serve as an indirect safety measure when accident data is scarce or unavailable, and these incidents are recorded by in-vehicle or street cameras (Guo et al., 2010). Despite the recognized importance of studying near-crash events, the relationship between such events and actual crashes remains uncertain and is being researched. According to some researchers, there is little correlation between crash rates and crash surrogates, such as traffic conflicts

(Dingus et al., 2016; Guo et al., 2010). They also pointed out that estimations of risks are conservatively derived from near-crash events. Crash data should be utilized in place of crash surrogates whenever readily available (Dingus et al., 2016; Guo et al., 2010). However, according to another study, crash surrogates are useful in advanced crash risk assessments and help distinguish between critical elements of successful crash avoidance maneuvers when crash data are scarce (Guo, 2019).

A study was conducted to determine if non-crash naturalistic driving data (NDD) could substitute for crash data in assessing the performance of automated emergency braking systems (AEB) compared rear-end crashes simulated from two different NDD sources with actual crash data. This revealed significant variances in impact speeds and scenario types. While NDD-based simulations effectively mimic pre-crash conditions, they significantly diverge from the real crash dynamics. This led to a recommendation for caution when using NDD to design and evaluate preventive safety systems, highlighting the simulations' limitations in accurately representing true crash conditions (Olleja et al., 2022).

Crash prediction is useful for taking corrective action on the causes of crashes through interventions using engineering, vehicle-assistive technologies, and behavioral interventions. Currently, surrogate indicators have become proactive approaches for the prediction of crashes. Various surrogate indicators are proposed as Time to Collision (TTC), Deceleration Rate to Avoid Crash (DRAC), Crash Potential Index (CPI), and Proportion of Stopping Distance (PSD) are applied in safety evaluation (Tarko et al., 2009).

The i-DREAMS technology also identified causes of safety breaches, which are contributing factors. A stakeholder survey was conducted to assess safety breaches and contributing issues on the i-DREAMS project to aid system development. The survey showed passenger cars, bus/coach, and truck modes of transport. The result revealed that the most critical safety breach for passenger cars is loss of control, and the contributing factor is excessive speeding. The most common safety breaches for bus/coach are loss of control and sudden brake, and the contributing factors are fatigue/sleepiness and inattention/distraction, respectively. However, Trucks' most crucial safety breach is closely following another vehicle; the contributing factors are stress, time pressure, and inattention or distraction. In general, loss of control, sudden braking, and close following another vehicle are the most common safety breaches identified in the survey, and inattention or distraction becomes the contributing factor (Talbot et al., 2020).

In order to understand and investigate the effectiveness of in-vehicle and post-trip interventions, many studies are conducted through simulators, instrumented cars, and self-reported studies. Among this, one study examined the effect of driver inattention on near-crash/crash risk using data from the 100-Car Naturalistic Driving Study. In-vehicle warning systems, including forward collision warning, lane departure warning, and curve speed warning, were beneficial in lowering the probability of rear-end and lane departure crashes (Klauer et al.,2006).

On the other hand, a study examined the impact of age, experience, and training on driving performance using a driving simulator. The study found that in-vehicle feedback and coaching systems, such as real-time speed, acceleration, and braking feedback, effectively improved driving performance and reduced risky driving behaviors among young and older drivers (Reimer et al., 2012).

More importantly, studies provided evidence to suggest that in-vehicle interventions can effectively reduce risky events in different risk zones. However, it is essential to note that the effectiveness of these interventions can vary depending on the specific intervention, the risk zone, and other contextual factors. Another study analyzed naturalistic driving data to evaluate the effectiveness of in-vehicle lane departure warning systems. The study found that these systems effectively reduced the frequency of lane departure events and increased the time to collision, particularly in high-speed and high-risk driving situations (Yan X. et al.,2017).

A study conducted a systematic review and meta-analysis of 26 studies on the effects of in-vehicle driver behavior feedback systems on risky driving behaviors showed that these interventions were effective in reducing risky driving behaviors, such as speeding and harsh acceleration, in various risk zones (Li, Y. et al.,2019). In-vehicle interventions, such as feedback devices, effectively reduce risky driving behavior. However, the effectiveness of these interventions can be influenced by the type of risky event being targeted. The effectiveness of in-vehicle and post-trip interventions can also vary depending on the location of the targeted risky events. For example, feedback on driving behavior is more effective in urban areas, where there are more opportunities for drivers to engage in risky behavior (Kircher et al., 2007). Similarly, roadside advertising signs have been found to have a more significant impact on driver behavior in rural areas, where there are fewer distractions on the road (Keall et al., 2004).

Furthermore, different investigations have been carried out to contrast the performance of post-trip and in-vehicle feedback interventions. The main conclusion was that providing real-time and post-trip feedback has a promising effect on enhancing driver performance. Dijksterhuis et al. (2015) came to a similar conclusion that post-trip feedback provides nearly identical benefits to real-time feedback. However, another study revealed that personalized and in-depth post-trip feedback is preferred over in-vehicle warning systems (Roberts et al., 2012).

2.4 Road infrastructure-related Variables and Risky Driving Events

Numerous studies have indicated a strong relationship between infrastructure characteristics and accident risk. Road infrastructure parameters significantly influence crash frequency, with various factors such as traffic volume, lane characteristics, and road geometry playing crucial roles. A recent study by Khattak et al. (2024) provides valuable insights into the relationship between road infrastructure parameters and crash frequency using Generalized Poisson (GP) models. The study found that crash frequency positively correlates with traffic volume ('Annual Average Daily Traffic' or AADT) and segment length, indicating that busier and longer roads are more prone to crashes. Additionally, the number of lanes was generally associated with reduced crash frequency, particularly in multi-vehicle crashes, suggesting that more lanes might ease congestion and improve traffic flow. However, this association was not significant for single-vehicle crashes. Lane width also played a crucial role, with wider lanes significantly reducing crash frequency across most crash types, likely by providing more space for safe maneuvering (Khattak et al., 2024).

Research on the impact of road infrastructure on traffic crashes highlights that road design elements significantly influence accident risk. Key aspects include road geometry, surface conditions, and roadside hazards. Highways with a relatively wider larger number of lanes generally have lower crash rates but higher severity, while arterial roads see higher crash frequencies due to traffic volumes, smaller number size of lanes, and frequent intersections (Kononov et al., 2008). A Dumbaugh and Rae (2009) study suggests that road design features, such as narrow lanes and sharp curves, contribute to higher crash rates on local roads. Thoughtful planning and adherence to safety standards in these areas can mitigate accident risks (Pembuain et al., 2019). This study employs a comprehensive literature review to link road infrastructure elements to safety outcomes, advocating for interventions such as Road Safety Audits (RSA) and Road Safety Inspections (RSI) to enhance road safety. Further evidence is provided by a comprehensive analysis of the safety impact of increased speed limits on rural highways in British Columbia, which found that higher speed limits led to a statistically significant increase in fatal and injury crashes by 11.1% (Kononov et al., 2008). This highlights the potential negative safety consequences of raising speed limits on rural roads, suggesting that higher speeds may result in more severe accidents.

Additional studies support these findings. For instance, Elvik (2013) conducted a meta-analysis on the effects of road design improvements and found that better road geometry, such as the presence of medians and improved road alignments, can reduce crash rates. Similarly, another study noted that roadside safety features, like guardrails and clear zones, are critical in preventing run-off-road crashes (Pembuain et al., 2019). Moreover, the safety performance of different pavement surfaces found that well-maintained road surfaces significantly reduce the likelihood of accidents, particularly in adverse weather conditions. This study underscores the importance of regular road maintenance and using high-quality materials to enhance road safety.

Overall, these studies collectively demonstrate the critical role of road infrastructure in influencing accident risk and highlight the complex relationship between road infrastructure parameters and crash frequency. This also indicated that different road categories are associated with varying levels of risk and types of accidents attributed to the varying geometric characteristics and levels of traffic interactions, underlining the importance of considering these factors in road design and safety interventions.

2.5 Driver-related variables and risky driving events

Recent studies have explored various factors contributing to driving risks, focusing on driver-related, vehicle-related, and infrastructure-related aspects. For example, using naturalistic driving data from 41 individuals, significant variables linked to near-crashes were identified, such as driver demographics and experience, including age, driving miles, and driving experience (Naji et al., 2018). Similarly, age, personality, and incident rates were critical in influencing crash events (Guo & Fang, 2013).

The impact of age and gender on road safety has shown mixed results. Drivers over 45 are at a higher risk of near-crashes (Naji et al., 2018). In contrast, younger drivers, particularly around the age of 26, are more prone to inattention-related crashes (Klauer et al., 2006). Another study indicated that drivers under 25 face the highest risk of crashes (Guo & Fang, 2013). Another study on crash risk by driver age, gender, and time of day found that drivers aged 30–39 had the highest rate of crashes during the day, while those aged 21–29 experienced the most crashes in the evening and nighttime. The frequency of crashes decreased with age, reaching its lowest point among drivers aged 70 and older. The study also noted that both men and women exhibited similar crash trends, though women had lower crash frequencies overall (Regev et al., 2018). Gender differences also show variability; higher incident rates among females in certain age groups have been observed, but this trend does not hold consistently across all data (Guo & Fang, 2013).

Additionally, the role of inattentive driving and secondary tasks in contributing to near-crashes and crashes has been emphasized, with significant percentages of such events involving distractions (Kong et al., 2021; Dingus et al., 2016).

Specifically, a study of the i-DREAMS trial experiments identified significant correlations between socio-demographic factors and safety-critical driving events. The study highlights that gender significantly influences acceleration, fatigue, and Vulnerable Road User (VRU) crash avoidance risks. Specifically, females exhibited a higher risk of acceleration events during in-vehicle interventions. At the same time, males were more prone to fatigue during web-based interventions and more likely to avoid VRU collisions during in-vehicle and smartphone-based interventions (Mohammed, 2022). Notably, variables like driving experience and education level showed no significant impact on risk. However, employment type did affect deceleration and distraction events, with laborers experiencing higher risks during in-vehicle and web-based interventions. Income levels also influenced acceleration events, with those earning between 4,000 and 5,000 euros monthly exhibiting more aggressive driving behaviors (Mohammed, 2022).

2.6 Relationship of Risky Driving Behavior and Traffic Crashes

Driving behavior is crucial in a driver's likelihood of being involved in crashes. Evaluating these behaviors through naturalistic driving data provides clear insights into a driver's involvement in potential crash scenarios. Various risky driving behaviors contribute to crash involvement, including tailgating, speeding, steering, acceleration, and deceleration.

A study to examine driver risky behavior and traffic crashes in Abu Dhabi revealed that close following and speeding significantly contributed to crashes, particularly in the city center (Alkaabi, 2023). Similarly, within the i-DREAMS project, another study identified tailgating and speeding as the most critical contributors to crashes (Alemu, 2024). The relationship between crashes and some of the risky driving events (tailgating, speeding, acceleration, and deceleration) has been studied and presented in the following sections.

2.6.1 Tailgating

Tailgating is driving dangerously close to another vehicle without maintaining a sufficient following distance, a common form of aggressive driving that often leads to accidents (Xu et al., 2021). A study found that approximately 62% of drivers reported experiencing aggressive tailgating, which frequently results in traffic incidents. The National Highway Traffic Safety Administration declared that rear-end crashes constituted one-third of the total crashes in 2015, where tailgating has been identified as one of the leading causes of rear-end crashes (Lee et al., 2002).

The severity of crashes due to tailgating is particularly pronounced in highway driving because the high speeds and close following distances exacerbate the impact (Carter et al., 1995).

2.6.2 Speeding

Speeding reduces reaction time, increases stopping distance, and often results in a loss of vehicle control, particularly on curves or wet roads. Speeding is a major risk factor for rear-end collisions and other accidents.

Charly and Mathew (2023) studied freeway road segments to identify driving performance measures through field measurements. They found that higher speed variability on curved road segments correlates with an increased frequency of crashes. Key factors such as speeding, jerk events, and mean gradients significantly contribute to estimating traffic crashes.

Several factors influence speed management, including road speed limits, traffic calming measures, geometric conditions, and driver behavior. Reckless drivers who exceed speed limits are more likely to be involved in crashes (Rahman et al. (2016)). Maintaining control over speed is crucial for preventing collisions, and distractions that lead to loss of speed control can increase the risk of accidents (Abd Rahman et al., 2020).

2.6.3 Acceleration and Deceleration

Acceleration and deceleration are commonly used to assess driving performance, with frequent changes in these behaviors often leading to severe crashes. Bagdadi (2013) compared jerk occurrences with historical crash data, suggesting that road segments requiring frequent braking might be associated with long-term unsafe conditions, potentially resulting in more crashes. A study measured bus driver acceleration consistency and its correlation with traffic crashes in Sweden, finding a significant link between erratic acceleration patterns and a higher incidence of accidents compared to smoother, more consistent driving styles (Bagdadi, 2013).

Pande et al. (2017) utilized GPS data from naturalistic driving to test the hypothesis that segments where drivers frequently apply hard braking might indicate crash-prone areas over time. Research aimed to link high-magnitude deceleration jerks observed in driving data with the crash frequency used linear referencing in ArcMap to integrate GPS data with road characteristics along US Highway 101 in San Luis Obispo, California. Negative binomial regression analysis was used, and the results showed a significant relationship between high-magnitude deceleration jerks and long-term crash frequency in these segments. The study also found that roadway curvature and auxiliary lanes did not significantly affect crash frequency in the highway sections considered.

2.7 Spatial Analysis

The field of road safety has increasingly leveraged Geographic Information Systems (GIS) as a crucial tool for advancing research and enhancing interventions. By analyzing the spatial patterns of traffic crashes and risky events, traffic safety specialists can pinpoint sections of roadways with high frequencies of incidents. Identifying accident hotspots and appending value-added data provides a more robust understanding of the underlying factors contributing to these high-risk areas. This enhanced understanding is essential for appropriately allocating resources for safety improvements, enabling targeted and effective interventions (Anderson et al., 2009).

2.7.1 Density and Hotspot Analysis

Density and hotspot analysis can identify high and low-risk zones within a study area. Density analysis calculates the density of features within a specified area, such as accident hotspots. It can apply various clustering methods like Natural Breaks (Jenks), Equal Interval, and Quantile classification to reveal significant patterns. This approach helps visualize areas with higher concentrations of incidents, clearly indicating where interventions may be needed.

Hotspot analysis, on the other hand, focuses on identifying areas with the highest and lowest concentrations of traffic crashes and risky events. Hotspot analysis allows researchers and policymakers to better understand the spatial distribution of risk by identifying significant clusters of high and low concentrations. This is crucial for implementing targeted safety measures and interventions to reduce accidents and enhance road safety. Hotspot analysis identifies areas where the frequency of incidents significantly deviates from the norm, highlighting critical zones that require attention (Alam & Tabassum, 2023).

2.8 Evaluation of Intervention Effectiveness

In the i-DREAMS project, two methodological approaches (before-after analysis and questionnaires) are proposed for evaluating the outcomes of the intervention (Katrakazas,2020c). Outcome evaluation, also called effect evaluation, applies to whether targeted factors changed as a result of the intervention or not. Questionaries are key indicators of qualitative measurements (i.e., performance objectives and change objectives), through which researchers can gain valuable information about key issues from a large proportion of drivers. Before-after analysis can be used for both quantitative (safety outcomes and safety promoting goals) and observed qualitative indicators (performance objectives, change objectives) (Katrakazas,2020c). Before-after studies are essential for evaluating the effectiveness of road safety interventions, with four common methods used to conduct these analyses. Each method addresses different aspects, such as regression to the mean (RTM), traffic volume changes, and temporal trends (Srinivasan et al., 2016).

Simple Before-After Study: This basic approach compares crash frequencies before and after intervention but does not adjust for RTM or account for changes in traffic volume or other temporal trends.

Before-After Study with Traffic Volume Correction: This method enhances the simple before-after study by adjusting for changes in traffic volume, typically using crash rates rather than crash counts. Although it helps align the data with traffic volume fluctuations, this method assumes a linear relationship between crash frequency and traffic volume, which is often disproven as the relationship can be nonlinear.

Before-After Study with Comparison Group: This approach incorporates data from an untreated comparison group to account for broader temporal effects that might affect crash data. It uses a comparison ratio to adjust the crash frequency at treatment sites, assuming similar trends between the treatment and comparison groups. This method does not address RTM unless treatment sites are selected based on their crash history.

Empirical-Bayes Before-After Study: The most comprehensive method, the Empirical-Bayes (EB) approach, adjusts for RTM, changes in traffic volume, and other temporal effects. It estimates the expected number of crashes if no intervention had occurred using Safety Performance Functions (SPFs) derived from a reference group similar in key risk factors to the treatment group. This method is robust in adjusting for variations in traffic volume and other site characteristics over time.

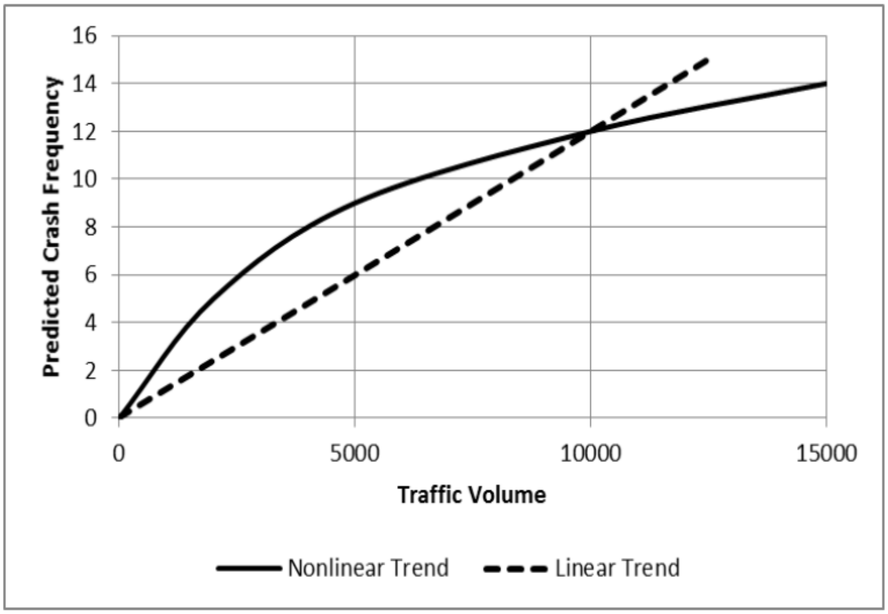

Figure 2.2: Relationships between crash frequency and traffic volume (Srinivasan et al., 2016)

2.9 Statistical Methods for Evaluation and Comparison of Interventions

Statistical methods can be used to model the relationships between variables and to evaluate the effectiveness of an intervention. Various naturalistic driving study designs employ different modeling methods to establish relationships between crash and near-crash risk factors. These designs can typically be categorized into frequency analysis, case-crossover, case-controlled, and responsibility study designs (Guo, 2019; Kim & Mooney, 2016). According to Kim and Mooney (2016), conflicting conclusions on risk factors often stem from biases inherent in these study designs. Recognizing the sources of these biases can lead to more informative future studies.

Various statistical methods are available to test the impact of interventions and compare their effectiveness across different groups. These methods are within the general before-after analysis framework and can employ one of the types under this framework presented in section 2.8. One commonly used method is Repeated Measures ANOVA, which tests whether three or more means are equal by examining the null hypothesis that the means of baseline and intervention phases are identical. Repeated Measures ANOVA assumes sphericity, which means equal variances of the differences between pairs of observation phases and requires the data to meet the assumptions of normality, homogeneity of variance, and independence of observations (Field, 2009). This method can determine if the rates of change between different groups are equal or if there are significant differences between groups. For example, Zhao & Wu (2013) used repeated measures ANOVA to analyze the effectiveness of a speed guidance system in a simulator study with 40 participants. Their results indicated that pre-warning and combining pre-warning and post-warning interventions could reduce overseeding. ANOVA can also efficiently compare responses across multiple groups and is widely used to examine demographic factors, such as age and gender, and assess driver behaviors with and without interventions (Field, 2009).

A more general approach is the Generalized Linear Model (GLM), where which includes various specific models and link functions, allowing researchers to account for non-normal data. The flexibility of GLMs makes them suitable for a wide range of data types and research questions. An extension to this is the Generalized Estimating Equations (GEE), which can further account for correlations within subjects due to repeated measures using a working correlation matrix. Zeger and Liang (1986) first used the GEE technique by extending the approach of a generalized linear model to correlated data in the context of repeated observations over time and acknowledged that GEE models generally are robust to misspecification of the correlation structure. Research by Mohammadi (2014) used a negative binomial model using a generalized estimating equation (GEE) procedure that incorporates the temporal correlations amongst yearly crash counts and indicated that the GEE model was effective in capturing temporal correlation and was found to be superior compared to the standard GLM.

The Poisson model is a method for analyzing count data. It assumes that events occur independently and at a constant rate over time or space. However, the Poisson model can be limiting if the data exhibit overdispersion, in which case the Negative Binomial Model would be more appropriate. The Poisson model is typically used for simpler count data scenarios where the mean and variance of the data are approximately equal (Cameron & Trivedi, 2013). For count data with overdispersion (where the variance exceeds the mean), the Negative Binomial Model (NBM) with a log link function is often preferred. This model adjusts for the overdispersion by introducing an additional parameter, making it more suitable for analyzing count data where events are not uniformly distributed (Hilbe, 2011).

Frequency models are the most common methods used for establishing relationships between variables, and they involve developing models for the frequency of crashes and crash surrogates (Guo, 2019). Poisson and Negative Binomial Regression are the most common statistical techniques for analyzing crash frequencies. In these regression analyses, the response variable is typically the frequency of crashes or crash surrogates, while the units of analysis are drivers or predetermined driving periods. Covariates may include driver characteristics such as age or frequency of kinematic events (Guo, 2019a).

For instance, Fitch et al. (2013) utilized a mixed-effect Poisson model to assess the risk of safety-critical events associated with mobile phone use. Guo et al. (2010) also employed this model to evaluate the frequency relationship between crashes and near-crashes, confirming a positive correlation between the two. Similarly, Ouimet et al. (2014) developed two longitudinal mixed-effect Poisson models to investigate the association between cortisol response and the rate of change in crash and near-crash (CNC) rates among teenage drivers.

Guo and Fang (2013) examined risk factors associated with individual driving risk using negative binomial regression models. They considered the frequency of CNC events as the response variable and miles traveled as exposure, with covariates including frequency of critical incidents, age group, gender, and personality score. Their study found that gender was not a significant variable, whereas the other factors had significant relationships with the frequency of CNC events. Cai et al. (2021) used Bayesian negative binomial models to demonstrate the association between various safety-critical events (SCEs) and crash outcomes, including crashes, injuries, and fatalities, among commercial truck drivers, providing robust evidence of a positive relationship between SCEs and injuries.

Multilevel modeling, also known as hierarchical, mixed-effect, random coefficient, or random parameter modeling, is another statistical method widely used to analyze crash frequency (Adanu et al., 2017; Adanu et al., 2019; Farmer et al., 2015; Freed et al., 2021). This method is suitable for studies with a hierarchical structure or multiple levels. Adanu et al. (2017) used multilevel models to evaluate the role of human factors in the variation of crash outcomes across different regions. They emphasized the importance of considering bidirectional effects between individuals and the social contexts or groups to which they belong. Most studies using multilevel modeling have focused on regional disparities affecting individual behaviors. However, Freed et al. (2021) demonstrated how individual driving performance could vary daily using multilevel models, showing the method's utility in personal behavior forecasting.

Despite its benefits, multilevel analysis is not widely utilized. Freed et al. (2021) noted their study was the first to analyze variability in distracted driving behavior using naturalistic driving data (NDD), finding that multilevel regression provided statistically sound results. Similarly, Dupont et al. (2013) utilized multilevel logistic regression to assess locality-related factors' impact on driver speeding behavior in NDS, highlighting the necessity of analyzing multiple observation levels for accurate road safety research conclusions.

2.10 Summary of Literature Review

Generally, there is a gap in the existing literature in integrating spatial and statistical analysis to effectively evaluate intervention impacts in various contexts such as risk zone, event severity, and influence of road infrastructure category. Combining spatial analysis with statistical modeling enhances the precision and depth of analysis, increasing the practicality of road safety research and supporting proactive interventions to protect human lives. Therefore, it is essential to consider both spatial and quantitative analyses in the context of local conditions, thereby contributing to the body of knowledge in this research domain.

A specific study within the i-DREAMS project aimed to predict crashes based on driver risk events identified tailgating and speeding as the most critical contributors to crashes (Alemu, 2024). Based on this fact, this study selected speeding and tailgating events for detailed investigation.

Even though a comprehensive analysis of the i-DREAMS trial experiments identified significant correlations between socio-demographic factors and safety-critical driving events, there was no significant difference in socio-demographic factors such as age, driving experience, gender, education, and income level in relation to speeding and tailgating events (Mohammed, 2022). These findings allowed this study to proceed without the need to control for demographic factors.

Studies suggest using comparison groups and considering traffic volume variations to evaluate intervention impacts effectively. This addresses different aspects, such as regression to the mean (RTM), traffic volume changes, and temporal trends (Srinivasan et al., 2016). However, the current study is based on secondary data from the i-DREAMS project, which was collected across the entire road network of Belgium over four phases where utilizing an untreated comparison group that would provide a valid benchmark against which to measure the intervention's effectiveness is inherently difficult due to the broad scope of the study. Similarly, traffic volume data, essential for adjusting the frequency of risky driving events to reflect varying levels of road use, was not available for the different phases of the intervention. Thus, this study employed the simple before-after analysis.

3. RESEARCH METHODS AND MATERIALS

3.1 Description of the Study Area

The framework of the i-DREAMS project, in general, is being tested using a simulator study and three stages of on-road trials in different European countries: Belgium, Germany, Greece, Portugal, and the UK. The test trial involves 600 drivers distributed to five countries and four Transport modes: cars, trucks, trains, and trams (Brown et al., 2023).

This study will be based on on-road trials of car vehicles in Belgium, and the study area has a land area coverage of 30.68 Km^2 and 11.5 million inhabitants. The population density of the region is four times higher than the average population density of the European Union.

3.2 Study Area Selection Criteria

The research study area was chosen based on the availability of naturalistic driving (baseline) data and intervention stage data in Belgium. Belgian road fatalities per million residents are slightly higher than the EU average of 42 deaths per million people; Belgium had a road fatality rate of 51 per million in 2020 (European Commission,2020). This reality required stakeholders and researchers to look into the underlying issues in order to provide practical intervention strategies to meet the targets of reducing the number of traffic fatalities to 30 per million people by 2030 and to zero by 2050. Belgium is one of the five European countries where the i-DREAMS on-road trials are being tested, and it has been selected for this study to make the scope more practical for an in-depth analysis of the effectiveness of the interventions.

3.3 Driving Risky Events

The driving risky events, which are the measurements of the change objectives in the i-DREAMS intervention framework, include Speeding, Acceleration, Deceleration, Tailgating, Steering, Vehicle control, Lane discipline, Overtaking, Road sharing, Fatigue, Distraction, Health, Forward collision avoidance and Vulnerable road user avoidance (Brijs, 2020). Some events have high, medium, and low levels, while others do not. A specific study within the i-DREAMS project aimed to predict crashes based on driver risk events identified tailgating and speeding as the most critical contributors to crashes (Alemu, 2024). The study employed a Negative Binomial model developed by using acceleration, deceleration, speeding, steering, and tailgating events as predictors and traffic crash count as response variables. Accordingly, the study determined that speeding, tailgating, and acceleration events have a significant and positive relation with the expected traffic crash, while steering and deceleration events are insignificant to the model to predict the expected traffic crash (Alemu, 2024). Based on this fact, this study selected speeding and tailgating events for detailed investigation due to their significant impact on crashes.

3.4 Data Collection

The researcher conducted the study based on i-DREAMS data on road trials for car drivers in Belgium. As explained above, i-DREAMS is a European Union project that aims to improve safe transportation. The three on-road trial stages of the i-DREAMS are detailed in Table 3.1. In this study, the observations in the third and fourth stages are to be used. The third stage (Stage 3) is the baseline phase without interventions; the fourth stage (Stage 4) is in-vehicle intervention and two post-trip interventions. All the drivers in the third stage participated in the fourth stage (Haancox, 2020).

Table 3.1: Stages of Data Collection on Simulator and On-road Trials (Hancox, 2020)

Stages	**Description**	**Duration**	**Remark**
Stage 1	Simulator Testing		(Not valid for the study)
Stage 2	Pilot Testing	Four weeks	(Not valid for the study)
Stage 3	Baseline (no intervention) Testing	Four weeks	Phase 1
Stage 4	In-vehicle Intervention	Four weeks	Phase 2
	Post-trip feedback on the smartphone	Four weeks	Phase 3
	Post-trip feedback on smartphones and gamified web platforms	Six weeks	Phase 4

The i-DREAMS project collected data on fourteen dependent variables, and their definition is taken from deliverable 3.3 of the i-DREAMS projects, a toolbox of recommended interventions (Brijs, 2020).

Definitions of Driver Risk Events:

1. Speeding: Total number of events where a driver exceeds the speed limit.
2. Acceleration: Total number of harsh and aggressive acceleration events.
3. Deceleration: Total number of harsh and aggressive deceleration events.
4. Steering: Total number of harsh cornering events.
5. Vehicle Control: Summation of acceleration, deceleration, and steering events.
6. Tailgating: Total number of headways monitor warnings.
7. Lane Discipline: Total number of lane departure warnings.
8. Overtaking: Total number of illegal overtaking events.
9. Forward Collision Avoidance: Total number of forward-collision warnings.
10. Vulnerable Road User Collision Avoidance: Total number of pedestrian collision warning events.
11. Road Sharing: Summation of tailgating, lane discipline, overtaking, forward collision avoidance, and vulnerable road user collision avoidance.
12. Fatigue: Total number of events caused by long hours of driving and driving at night.

13. Distraction: Total number of events involving handheld phone use and hands not on the wheel.
14. Health: Summation of fatigue and distraction events.

3.5 Data Manipulation

The original dataset from the i-DREAMS project was cleaned, extracted, and manipulated through the latest version of GIS software and Excel to review the location, distribution, date, type of events, and severity, mainly focusing on Speeding and Tailgating events. This study used speeding and tailgating data on 52 drivers in Belgium.

Table 3.1: Extracted Tailgating and Speeding Events across the 4 phases

	Phases			
Events	Phase 1	Phase 2	Phase 3	Phase 4
Tailgating (count)	32,342	31,005	32,721	40,933
Speeding (count)	9,931	9,982	9,640	13,248

3.6 Data Analysis

The analysis in this study consisted of two parts: Spatial Analysis and Statistical Analysis. The Spatial Analysis uses speeding and tailgating events baseline phase data (naturalistic driving conditions) to identify high-risk (hotspots) and low-risk zones within the study area. Further, the spatial analysis joined risky driving events with different classes of the road network, enabling the filtering of events along these road categories across different phases for further comparison of the intervention among these different categories. The Statistical Analysis aimed to evaluate the significance of intervention effectiveness and compare the effectiveness of interventions in high and low-risk zones and across different roadway categories. Based on the spatial analysis performed on Q-GIS, the statistical analysis provided insights into the impact of interventions and identified areas and risky driving events where they are more effective.

3.6.1 Spatial Analysis

Spatial analysis is crucial in road safety research, particularly in identifying risky zones and hotspot sections of roads. Identifying hotspots, crash locations, and areas with high crash concentrations is a common task for road safety organizations and researchers (Zandi et al., 2023). The introduction of spatial autocorrelation has significantly enhanced the ability to model relationships between variables in space through clustering. This method creates associations between different variables, leading to more scalable outcomes for road safety research. For instance, a study by Yang et al. (2021) proposed a new safety performance measure called Risk Status (RS) by fusing crash data and surrogate safety measures, which in turn enabled the assessment of traffic safety.

Gomes et al. (2017) evaluated the performance of GWR in a Generalized Linear Model (GLM) context. They found that accounting for the over-dispersion of crash data necessitated using Geographically Weighted Negative Binomial Regression (GWNBR) for spatial analysis of traffic crashes. Furthermore, Xu and Huang (2015) extended GWR to a semiparametric model by combining geographically varying parameters with geographically constant parameters.

This study conducted spatial analysis on speeding and tailgating events using baseline phase data (naturalistic driving conditions). The primary aim was to identify high-risk (hotspots) and low-risk zones within the study area. The spatial analysis also involved spatially joining risky driving events with different classes of the road network, allowing for the filtering of events along these road categories across different phases. This approach facilitated relevant comparisons between different groups and phases to address the study's research questions effectively.

3.6.1.1 Identification of High and Low-Risk Zones

Density mapping of the two risky events was done to identify the significant hot and cold spots of the study area, which is a crucial step to compare the effect of the different intervention stages of the i-DREAMS project among the high and low-risk zone risky events. The spatial distribution of risky events observed through the latest Q-GIS software also gives a clear insight into understanding the risky events' hotspot locations. This analysis relies on density mapping of the risky events during the baseline phase, employing the Natural Breaks (Jenks) optimization method. The Natural Breaks (Jenks) classification is a data clustering technique that finds the best way to divide values into distinct classes (Ke et al., 2023). It works by minimizing the variance within each class while maximizing the variance between classes, thereby optimizing the uniqueness of each class. This method is particularly effective for highlighting natural groupings inherent to the data, making it an ideal choice for this type of spatial analysis.

3.6.1.2 Spatial Analysis of Risky Driving Events by Road Class

The second analysis in Q-GIS involved spatially joining risky driving events with various classes of the road network. This process enabled the filtering of events along different road categories across multiple phases. Initially, motorways and secondary road features from OpenStreetMap were imported to create road polygons. These polygons were then used to spatially join with the risky events data, selecting the events along each road class by finding their intersections. The resulting filtered data was subsequently manipulated and utilized for statistical analysis to assess the effectiveness of interventions across these different road categories.

3.6.2 Statistical Analysis

Statistical analysis is crucial in road safety research, where various models are developed using conventional statistical methods and machine learning techniques. In the context of risky driving events, it is essential to consider the spatial and temporal variability of the data, which necessitates localized analysis and predictive models. The factors contributing to risky driving events vary based on the environment, road geometry, driver condition, and vehicle factors. Therefore, models that can check the significance of intervention considering this variation should be employed.

In this study, the primary focus of the statistical analysis is not to develop predictive models but to evaluate the significance of intervention effectiveness. This involves comparing the effectiveness of interventions in high and low-risk zones and across different roadway categories based on the spatial analysis performed by Q-GIS to provide insights into the impact of interventions and identify areas where they are most effective.

The chosen statistical methods allowed for rigorous hypothesis testing, enabling the study to assess whether the observed changes in driving behavior were statistically significant. This involved comparing the frequency and severity of events across different phases of the intervention while also considering the variations across high-risk and low-risk zones and different road categories (e.g., motorways, primary roads, and secondary roads).

3.6.2.1 Characteristics of the Data and Statistical Method Selection

The dataset for risky driving events in this research exhibits several distinctive characteristics:

1. Count Data with Overdispersion

- The data represents count occurrences of risky driving events, such as speeding and tailgating. The data exhibit overdispersion, where the variance exceeds the mean, indicating more significant variability than what is assumed by standard Poisson models.

2. Non-Normal Distribution

- The distribution of both speeding and tailgating data is non-normal, exhibiting high skewness, which violates the assumptions of many parametric statistical tests, such as the Repeated Measure of ANOVA.

3. Correlated Data Due to Repeated Measures

- The data includes repeated measures for the same subjects (drivers) across different intervention phases. This introduces correlation within the data, as measurements from the same subject are likely to be more similar to each other than to those from different subjects.

4. Unequal Number of Participants in Different Groups

- The number of participants varies across different groups, such as risk zones, as some groups from the spatial analysis did not have all drivers involved. Unequal group sizes can complicate the analysis and interpretation of results, mainly when using methods that assume equal sample sizes.

Correlation between individual data points is introduced when there is a hierarchical structure or repeated measurements exist. GLMM (Generalized Linear Mixed Models) and Generalized Estimating Equations (GEEs) model the data, taking into account this correlation (Vagenas et al., 2018). The two approaches differ in how they handle this correlation, and selecting between them relies on the research aims and study characteristics. GEE uses the working correlation matrix that helps to correct correlations due to repeated measures and provides estimates of the average population effect rather than subject-specific effects. GLMM, on the other hand, uses random effects to explicitly model the correlations within clusters or repeated measures (Vagenas et al., 2018). GLMM requires the distribution of the random effects and is more computationally intensive than GEE. Additionally, the focus of this study is not on the individual subjects or drivers but rather on the difference between different groups (clusters) and phases.

Accordingly, this study employed Generalized Estimating Equations (GEE) under the Generalized linear model framework but specifically required a working correlation matrix that accounts for repeated measures. A Negative Binomial distribution and a log link function with a robust correlation matrix are used to address these issues. This method is well-suited for the following reasons:

- *Handling Overdispersion*: The Negative Binomial model is specifically designed to handle overdispersion in count data, making it more appropriate than the Poisson model.
- *Non-Normal Data*: GEEs do not assume the normality of the response variable. Instead, they model the mean of the response variable as a function of the predictors through a link function.
- *Correlated Data*: Using a GEE allows for the inclusion of repeated measures and the accommodating of the correlation within the data due to repeated measures.
- *Unequal Group Sizes*: GEEs can handle unequal sample sizes across groups, providing robust estimates without requiring equal group sizes.

The Negative Binomial Regression model is used to model count data, where the response variable Y follows a negative binomial distribution. The mean μ of Y is related to the predictor variables $X1$, $X2$,…. Xk, through a log link function:

$$log(\mu) = \beta 0 + \beta 1X1 + \beta 2X2 + \cdots + \beta kXk \text{………………………………} \textit{Equation 3-1}$$

The variance of Y in a Negative Binomial model is given by:

$\mathrm{Var(Y)} = \mu + \alpha\mu^2$, Where α is the dispersion parameter.

The GEE extends this model by accounting for the correlation within clusters or repeated measures. The estimating equation for the parameters β *is*

$$\mathbf{S}(\beta) = \sum_{i=1}^{N} \mathbf{D}_i'[\mathbf{V}_i(\alpha)]^{-1}(\mathbf{Y}_i - \mu_i) = 0$$

where:

- $\mathbf{D}_i$ is an $n_i \times p$ matrix with (i, j)th elements $\frac{\partial \mu_{ij}}{\partial \beta}$.
- $\mathbf{V}_i(\alpha)$ is the working variance-covariance matrix of $\mathbf{Y}_i$.
- $\mathbf{Y}_i$ and μ_i are n_i-vectors with elements y_{ij} and μ_{ij}.

The Wald chi-square test is a statistical test used to evaluate the significance of individual predictors or variables within the Negative Binomial Regression model. The test assesses whether the estimated coefficients of the predictors are significantly different from zero, indicating that the predictors have a meaningful impact on the response variable (Park et al.,2013).

For each coefficient βj, the Wald statistic is calculated to test the null hypothesis $H0$: βj=0 against the alternative hypothesis $H1$: βj≠0. The Wald statistic is given by:

$$Wj = \frac{\hat{\beta} j^2}{Var(\hat{\beta} j)} \text{ ……………………………………………………………… } \textit{Equation 3-2}$$

Whare $\mathrm{Var}(\hat{\beta}\mathrm{j})$ represents the robust variance estimate of the coefficient $\hat{\beta}\mathrm{j}$ which is adjusted based on the working correlation matrix in GEE.

4. RESULTS AND DISCUSSIONS

4.1 Spatial Analysis

Spatial analysis of tailgating and speeding events was conducted using the QGIS software package. This included mapping the distribution of these events within the study area, identifying density-based clusters and hotspot areas, spatially joining events with various roadway categories, and performing a range of spatial analyses on the data, such as identifying high and low-risk zones through clustering and spatial joining of events with different road types.

4.1.1 Mapping of Tailgating and Speeding Events

The data on tailgating and speeding events was filtered and processed using the latest Q GIS software to visualize their distribution. Based on baseline (no intervention) data from 52 car drivers in Belgium, the resulting map shows that these events are predominantly concentrated in major cities and roads with higher traffic interaction. The distribution also reveals that most drivers were traversing the Limburg province in the Flanders region of Belgium.

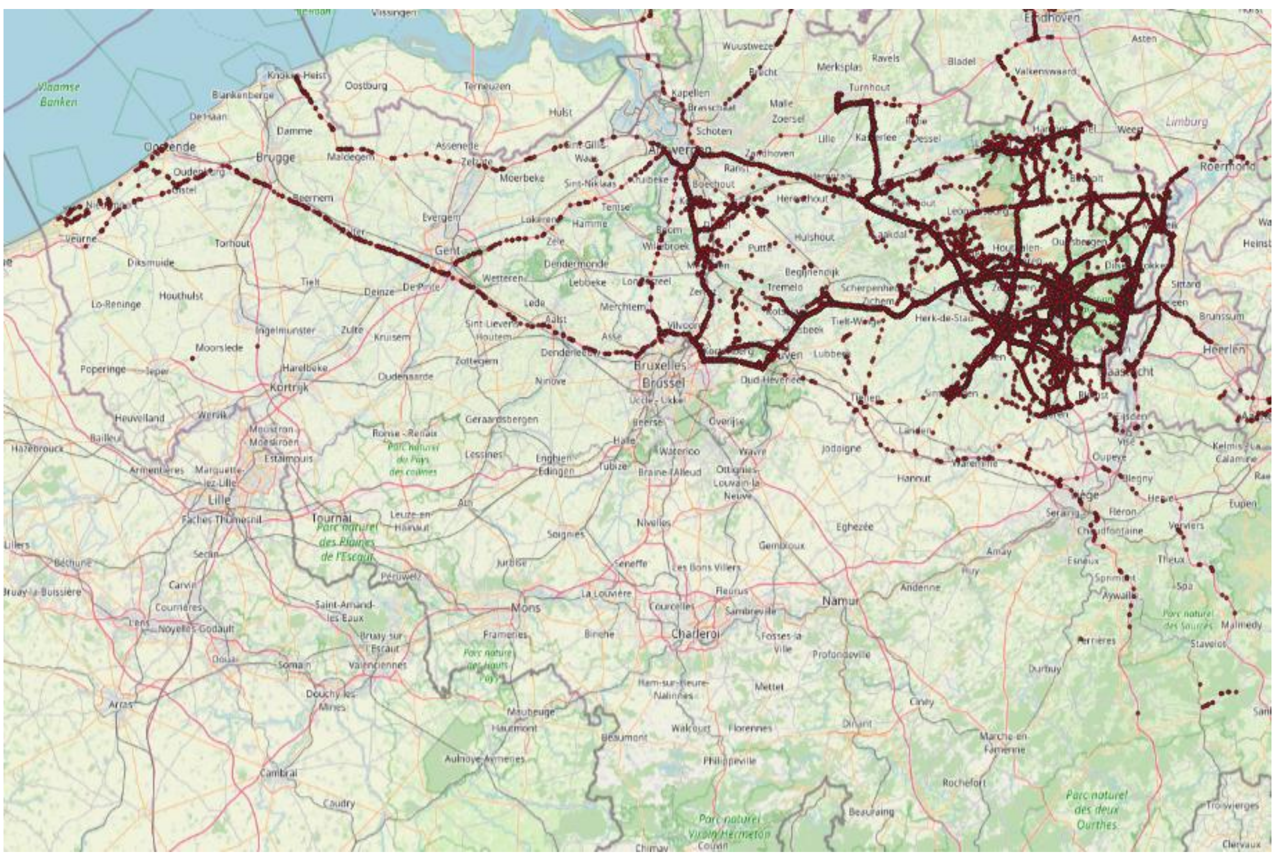

Figure 4.1: Distribution of Tailgating Events in Study Area

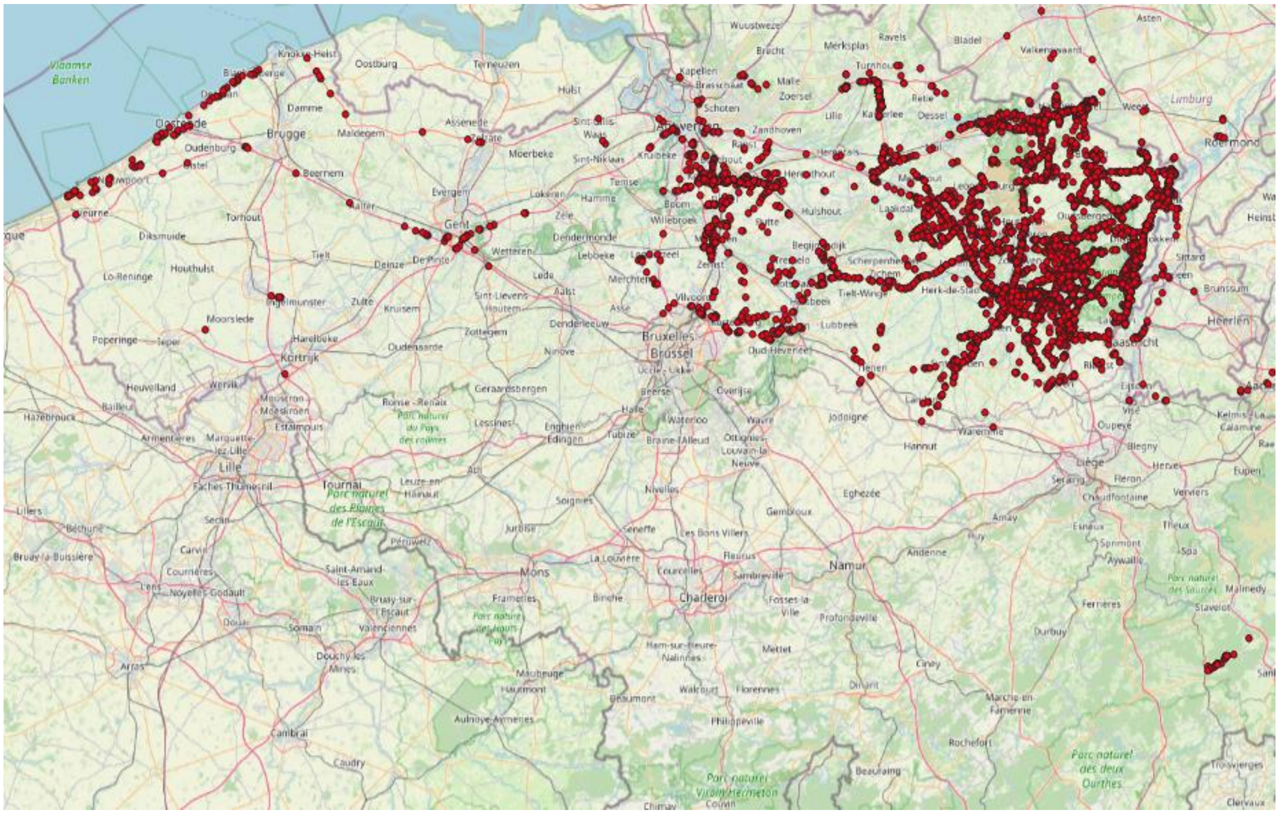

Figure 4.2: Distribution of Speeding Events in Study Area

4.1.2 Mapping and Clustering

Clustering involves dividing a dataset into groups, or clusters, where data points within the same group are as similar as possible to each other, while data points in different groups are as distinct from each other (Aliguliyev, 2009). The study employed spatial analysis on speeding and tailgating events using baseline phase data collected under naturalistic driving conditions. Density clustering, particularly using the Natural Breaks (Jenks) optimization method, was a key component of this analysis. The Natural Breaks (Jenks) optimization method is specifically designed to enhance the clarity of data visualization by grouping data into classes most naturally associated with each other (Ke et al., 2023; Nyimbili et al., 2020). This method works by minimizing the variance within each class while maximizing the variance between classes, thus creating a clear separation between different groups. In practical terms, this means that the data is divided into classes, and each class contains values that are similar to each other and distinct from those in other classes.

This study applied the Natural Breaks (Jenks) method to the spatial distribution of speeding and tailgating events. The analysis identified natural groupings within the data, corresponding to areas of varying risk levels. This technique was particularly effective in highlighting high-risk zones, areas where the density of risky driving events was higher than in other regions. These high-risk zones, or hotspots, were primarily located in major cities and high-traffic areas within the study area.

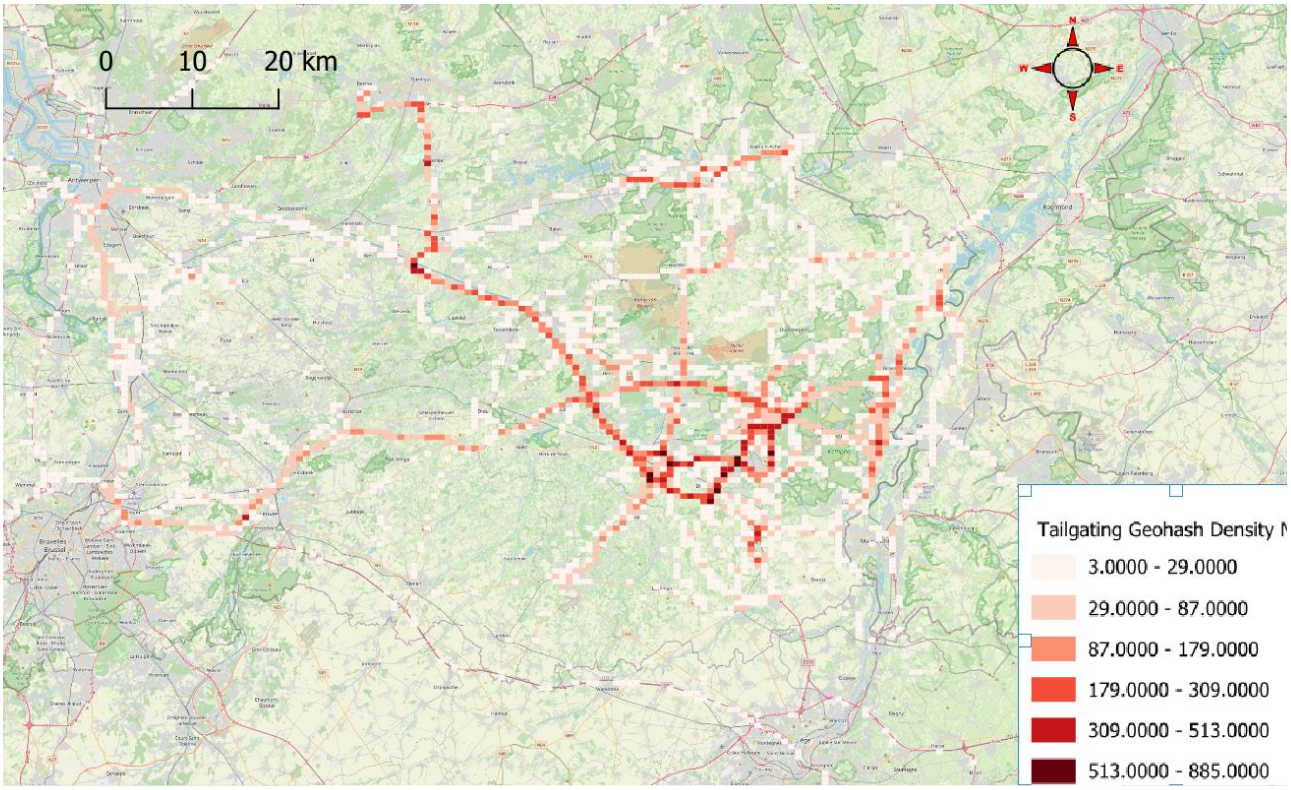

Figure 4.3: Density-based clustering for Tailgating Events

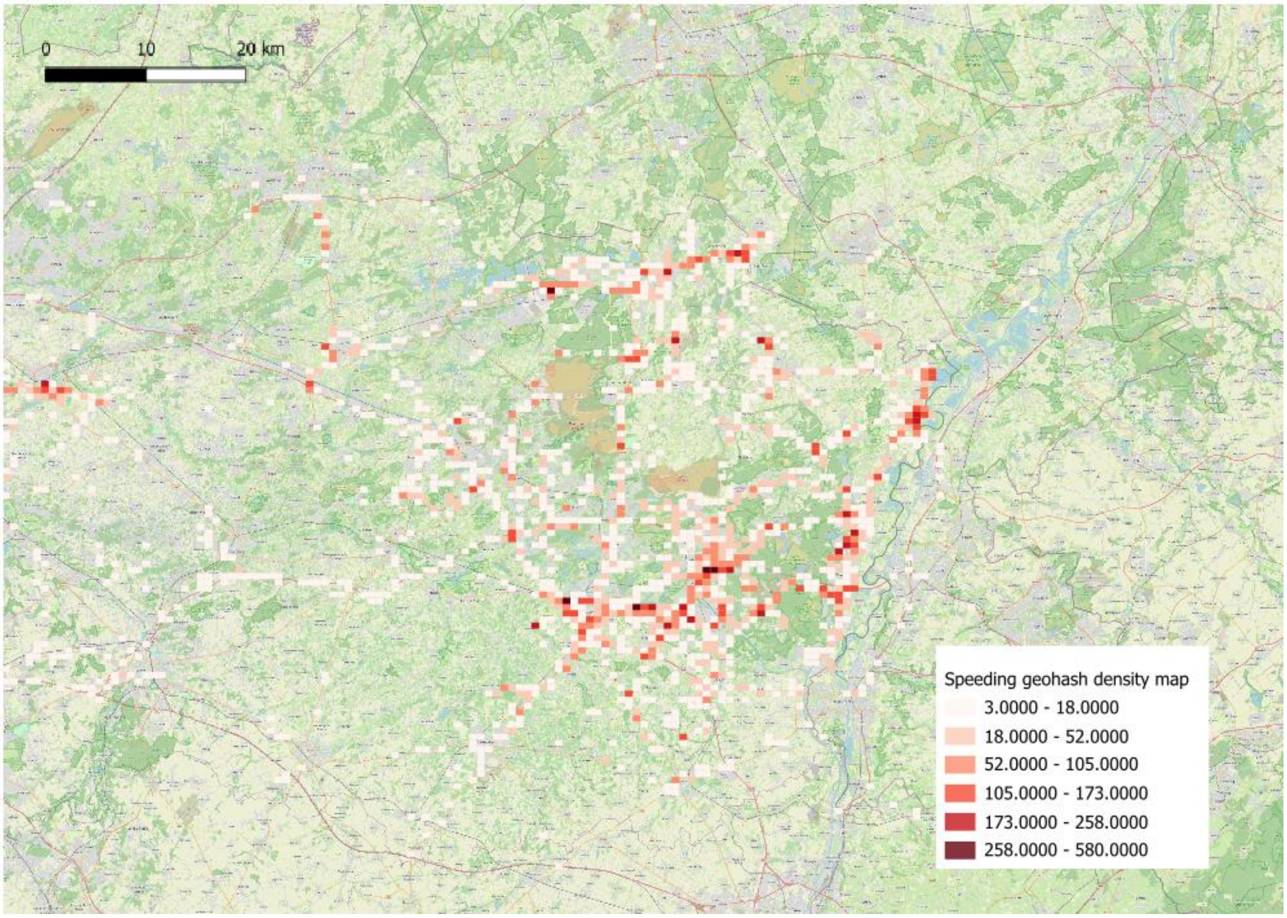

Figure 4.4: Density-based Clustering for Speeding Events

4.1.2.1 Spatial Analysis by Risk Zone

In the context of this study, spatial analysis was a fundamental tool for understanding the distribution and patterns of risky driving behaviors, specifically speeding and tailgating events. The analysis was based on baseline phase data collected under naturalistic driving conditions. As illustrated by the distribution of tailgating and speeding events in the study area, it is easy to identify the areas with the highest number of these risky driving events.

The density mapping was utilized to closely identify zones with significantly higher risks (hot spots) and lower risk zones. This step was essential for comparing the effects of different intervention stages of the i-DREAMS project on risky driving events in high and low-risk zones. Utilizing the Q-GIS software, the spatial distribution of risky events was analyzed, providing clear insights into the locations of these hot spots. This process allowed for the visualization of the speeding and tailgating incidents that were most concentrated. The following maps demonstrate events found in these high and low-risk zones in detail.

Figure 4.5: Tailgating Events in Hot-Spot Areas (highlighted in yellow)

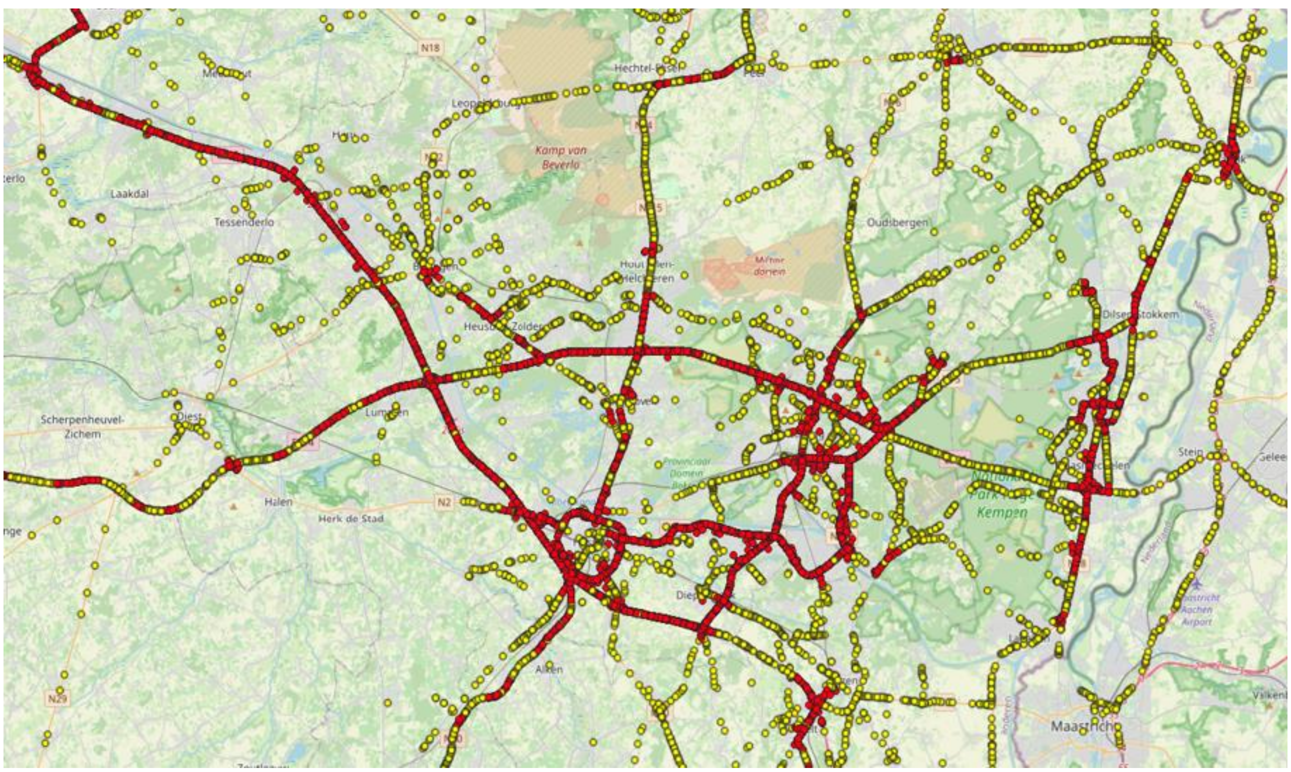

Figure 4.6: Tailgating Events in a Low-risk Zone (highlighted in yellow)

Figure 4.7: Speeding Events in High and Low-risk Zones

4.1.3 Spatial Joining of Events with Different Roadway Categories

The spatial analysis extended to spatially joining risky driving events with various classes of the road network. This process was crucial for filtering events along different road categories across multiple phases. Initially, road features such as motorways and secondary roads were imported from OpenStreetMap to create road polygons. These polygons were then used to spatially join with the risky events data, enabling the selection of events along each road class by identifying their intersections. These events were then filtered along motorways, primary roads, and secondary roads for further statistical analysis and to compare the effectiveness of interventions.

Table 4.1: Summary of Tailgating and Speeding Events count in different categories based on Spatial Analysis

Tailgating Events Count					
Categories		Phase 1	Phase 2	Phase 3	Phase 4
By Severity	High-Severity	5749	5201	5572	4611
	Medium-Severity	26593	25804	27149	22680
By Risk Zone	High-Risk Zone	14936	5154	5034	4299
	Low-Risk Zone	13613	11154	12255	9251
By Road Types (High-Risk Zone)	Motorways	2742	939	693	595
	Primary Roads	4030	3683	3347	2211
	Secondary Roads	626	687	616	404
	Motorways	3427	3264	3723	2519
	Primary Roads	3686	3227	3007	2131

By Road Types (Low-Risk Zone)	Secondary Roads	2168	1951	1756	1279
Speeding Events Count					
Categories		Phase 1	Phase 2	Phase 3	Phase 4
By Severity	High-Severity	7022	7134	6883	6155
	Medium-Severity	2909	2848	2757	2677
By Risk Zone	High-Risk Zone	1361	1156	974	897
	Low-Risk Zone	4684	3907	3859	3429
By Road Types	Motorways	652	551	631	539
	Primary Roads	2957	2754	2723	2466
	Secondary Roads	1677	1744	1790	1394

4.2 Descriptives Based on Event Counts

4.2.1 Event Count Based on Severity

Analyzing speeding and tailgating events based on severity levels provided key insights into the nature of risky driving behaviors observed in this study. The figures show a distinct pattern where most speeding events are categorized as high-severity while tailgating events predominantly fall under the medium-severity category. This result can be related to many factors, one of which can be the threshold limits defined for the high and medium-severity speeding and tailgating events. On the other hand, this finding can be attributed to the tendency of the participants to engage in high-severity speeding events, which is consistent with previous studies within the i-DREAMS project, which reported that Belgian drivers exhibit notably more risk-taking behavior than drivers in Germany and the UK. In particular, a significant proportion of Belgian drivers admitted to frequently exceeding the speed limit, with 94% acknowledging such behavior when asked (Brown et al., 2023).

This high incidence of severe speeding events suggested that Belgian drivers are more prone to engage in speeding, which poses a significant challenge for road safety interventions. This study also proved that the i-DREAMS intervention struggled to produce statistically significant improvements in reducing high-severity speeding events while aiming to mitigate such risky behaviors.

On the other hand, tailgating events were predominantly of medium severity, which might explain why the study observed more encouraging results in this area. Further statistical analysis performed in section 4.4 of this study found that the intervention's impact was more noticeable in reducing tailgating incidents, particularly in high-risk zones, where statistically significant reductions were recorded. This difference in intervention effectiveness between speeding and tailgating might be attributed to the lower severity of tailgating events, making drivers more receptive to corrective measures provided by the i-DREAMS system.

Generally, the severity-based analysis highlighted a clear contrast between speeding and tailgating behaviors, with the former being more resistant to intervention. At the same time, the latter showed more promising results in response to the i-DREAMS system, particularly in high-risk zones.

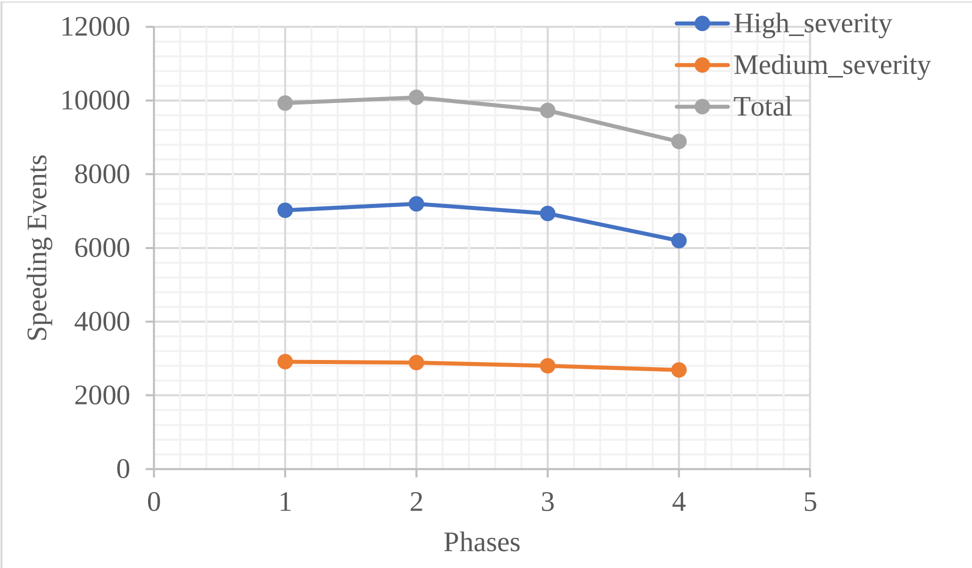

Figure 4.8: Speeding event counts based on event severity group

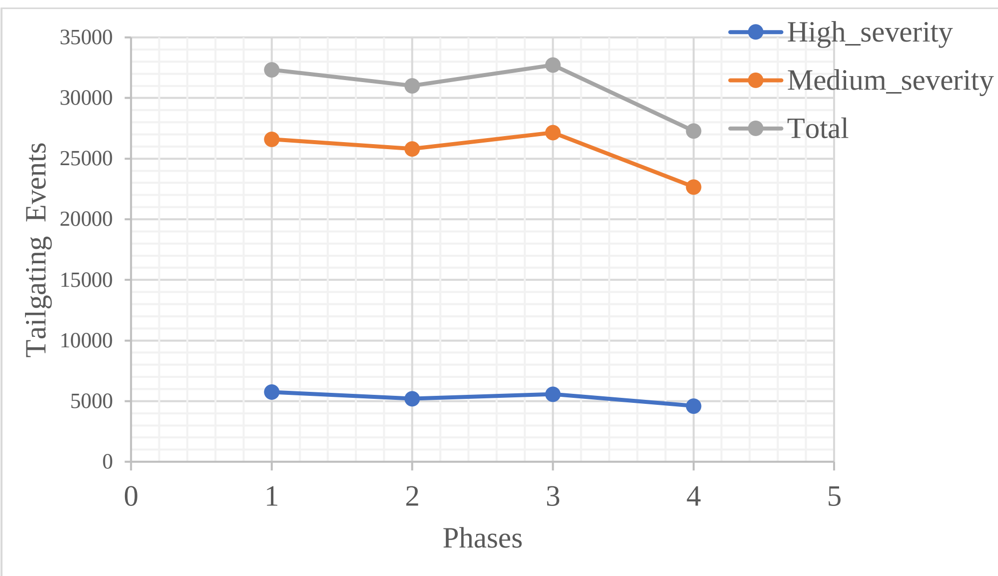

Figure 4.9: Tailgating event counts based on event severity group

4.3 Statistical Analysis

The count data of speeding and tailgating incidents within the study area were carefully filtered and processed using the latest GIS software to prepare it for detailed statistical analysis and comparisons. This preprocessing step ensured the data was accurately geolocated and categorized, allowing for precise spatial and temporal analysis of risky driving behaviors. The subsequent statistical analyses aimed to evaluate the effectiveness of the i-DREAMS intervention across different phases, considering the repeated measures of driving events.

Given the nature of the data and the objectives of the study, several statistical methods were considered to assess the impact of the intervention. The choice of the appropriate statistical model depends heavily on the characteristics of the data, such as whether the data meets the assumptions of normality, the presence of repeated measures, the count nature of the data, and whether there is overdispersion (i.e., variance greater than the mean). Considering these issues, this study employed Generalized Estimating Equations (GEE) with a working correlation matrix to handle the correlation due to repeated measures. The analysis used a negative binomial model with a log link function to analyze the count data of speeding and tailgating events. This model was selected because it effectively handles overdispersion, non-normality, and the unequal number of participants across different groups, which are common issues in data count from naturalistic driving studies. Additionally, the model's capacity to accommodate repeated measures within the data made it well-suited for this study's evaluation of the intervention's impact over different phases. The latest version of IBM SPSS was used to perform the statistical analyses in this study. Andy Field's book (2009) was consulted in testing the assumptions and performing and reporting the statistical analyses.

4.3.1 Test of Assumptions of the Method

The Generalized Estimating Equations (GEE) with a Negative Binomial distribution has several key assumptions that must be met for the model to be appropriate and the results to be valid. These assumptions include (Andy Field, 2009).

1. The dependent variable needs to be count data or discrete variables: The dependent variables in this study are the count of speeding and tailgating events.

2. The data needs to show overdispersion, which means the variance of the data is larger than the mean: The data for the speeding and tailgating events exhibited over-dispersion as presented in Table 4.2; this nature of the data also necessitates a negative binomial model. The unit of analysis is the count of High and Medium-severity speeding and tailgating events in each phase of the intervention.

Table 4.2: Test of overdispersion for speeding and tailgating events

		Speeding		Tailgating		Overdispersion	
Group	Phase	Mean	Variance	Mean	Variance	Speeding	Tailgating
High-Severity	1	135.06	17615.54	119.77	57851.68	Yes	Yes
Medium-Severity	1	56	2726.49	519.72	368022.6	Yes	Yes
Total	1	95.53	11657.19	331.51	260295.59	Yes	Yes
High-Severity	2	138.42	13972.81	108.35	20080.44	Yes	Yes
Medium-Severity	2	55.48	1823.95	501.56	190877.77	Yes	Yes

Total	2	96.95	9559.33	316.52	148405.45	Yes	Yes
High-Severity	3	133.38	11288.39	116.08	28114.5	Yes	Yes
Medium-Severity	3	53.81	1653.82	551.76	273946.76	Yes	Yes
Total	3	93.6	8019.74	346.74	204594.78	Yes	Yes
High-Severity	4	74.15	4565.42	96.06	12185.24	Yes	Yes
Medium-Severity	4	51.71	2039.34	438.46	143129.17	Yes	Yes
Total	4	62.93	3397.38	277.33	110274.75	Yes	Yes

3. Independence of Observations between subjects: The observations between subjects should be independent. However, for repeated or clustered measures, the model can account for correlations within subjects (clusters) or repeated measures using the working correlation matrix.

4. No Perfect Multicollinearity: No perfect multicollinearity among the independent variables should exist. Multicollinearity can inflate the variance of the coefficient estimates and make the model unstable. Thus, the correlation between the predictors (phases) should be checked, and those with high correlations must be avoided in the model. However, the predictors in this case (phases) are categorical variables; multicollinearity is less of a concern.

A correlation matrix may check multicollinearity, and when computing a matrix of Spearman's rank correlation among all factor variables, the correlation magnitude should be less than 0.8.

Table 4.3: Correlation coefficients between predictors (phases) of Tailgating and Speeding events

Correlation Coefficients (CC) between phases for Tailgating events						
			Phase 1	Phase 2	Phase 3	Phase 4
Spearman's rho	Phase 1	CC	1.000			
	Phase 2	CC	0.680	1.000		
	Phase 3	CC	0.638	0.783	1.000	
	Phase 4	CC	0.431	0.412	0.524	1.000
Correlation Coefficients (CC) between phases for Speeding events						
			Phase 1	Phase 2	Phase 3	Phase 4
Spearman's rho	Phase 1	CC	1.000			
	Phase 2	CC	0.791	1.000		
	Phase 3	CC	0.776	0.819	1.000	
	Phase 4	CC	0.262	0.173	0.279	1.000

As shown in Table 4.3, the correlation matrix for the speeding and tailgating events showed that the Correlation Coefficients (CC) are less than 0.8, except for a slight deviation between phase 2 and phase 3 speeding events.

5. The relationship between the independent variables and the dependent variable should be correctly specified: The model uses a negative binomial model with a log link function, which implies that the relationship between the independent variables and the log of the expected value of the dependent variable is linear.

4.4 Generalized Estimating Equations (GEE)

Generalized Estimating Equations (GEE) are used with a robust correlation matrix to handle the correlation due to repeated measures. The robust estimator is used as the covariance matrix as it has the property of being a consistent estimator of the covariance matrix even if the working correlation matrix structure is wrongly specified (Liang et al., 1986). Autoregressive correlation structure is used as it is best suited for data where repeated measurements are equally spaced in time, and the correlation between measures decreases as the time or phases progress. A Negative Binomial Model with a log link function was employed to evaluate the intervention's effectiveness and compare results among different groups. The statistical analysis began by testing the general significance of the intervention on tailgating and speeding events, including an assessment of travel distance as a covariate. Building on the spatial analysis results, further examination compared the intervention's effectiveness on tailgating and speeding events between high and low-risk zones. Additionally, the analysis extended to compare the impact of the intervention across various road categories, providing a comprehensive understanding of its effectiveness in different contexts.

4.4.1 Checking Distance as a Covariate

The statistical analysis of the intervention's impact on tailgating and speeding events was conducted both with and without considering distance as a covariate. This approach aimed to determine whether the distance traveled significantly influenced the interventions' effectiveness across different phases. The findings revealed no significant interactions between distance and phases, indicating that distance did not play a critical role in altering the impact of the intervention phases. The total travel distance covered by the participants in each phase is summarized in Table 4.4 below.

Table 4.4: Travel distance across the phases

Phases	Phase 1	Phase 2	Phase 3	Phase 4
Distance (km)	47534.42	45039.12	48470.44	65926.66

The results demonstrated that the effect of the intervention phases on tailgating and speeding events remained statistically insignificant regardless of whether distance was included as a covariate, as presented in Table 4.5. While distance itself is a significant predictor of the number of risky events (indicating that more distance covered correlates with more risky events), this finding is relatively straightforward and expected. Our primary interest lies in the interaction term (phases* distance), which tests whether the effect of intervention phases on speeding and tailgating events changes depending on the distance covered in each phase. The interaction term phases*distance was found to be insignificant (Wald Chi-Square = 3.308, p = 0.386) for tailgating events and (Wald Chi-Square = 6.898, p = 0.075) for speeding events. Moreover, considering distance as a covariate in the model resulted in unexpectedly higher mean values for the intervention phases than the baseline phase, as presented in Table 4.6. This outcome was deemed unrealistic, indicating that accounting for distance did not reveal any meaningful or statistical differences between the phases when comparing the adjusted marginal means considering travel distance. Therefore, subsequent analyses were conducted without including distance as a covariate, adjusting the number of phase 4 events. Further details and results can be referred to in the SPSS outputs found in Appendix B.

Table 4.5: Tests of model effects with and without distance as a covariate

Covariate	Events	Source	Wald Chi-Square	Sig.
-	Speeding	Phase	0.420	0.936
	Tailgating	Phase	1.592	0.661
Distance	Speeding	Phase	6.711	0.082
		Distance	40.202	<0.001
		Phase*Distance	6.898	0.075
	Tailgating	Phase	1.874	0.599
		Distance	127.633	<0.001
		Phase*Distance	3.038	0.386

Table 4.6: Estimates of speeding and tailgating events with and without distance as a covariate

		Without Distance		**With Distance**	
Events	Phase	Mean	Std. Error	Mean	Std. Error
Speeding	Phase 1	194.73	24.517	164.78	18.208
	Phase 2	195.73	21.807	178.97	16.929
	Phase 3	189.02	20.000	167.90	15.924
	Phase 4	178.43	19.919	178.72	19.642
Tailgating	Phase 1	654.45	116.992	437.75	40.226
	Phase 2	587.65	72.962	482.96	34.460
	Phase 3	641.57	88.817	488.24	45.180
	Phase 4	523.24	66.672	471.49	55.511

4.4.2 Analysis by Event Severity

In this section, the analysis aimed to determine if there were significant differences between the baseline and intervention phases for tailgating and speeding events and to compare the rate of change of high-severity and medium-severity events across the phases for both types of risky driving behaviors. The summary of the descriptives by event count indicated that speeding events were predominantly classified as high severity, highlighting the persistent challenge in mitigating such risky behavior. In contrast, tailgating events in this study were mostly of medium severity, and the following section will statistically evaluate the changes in these behaviors across the different intervention phases.

4.4.2.1 Tailgating

Based on the statistical analysis results, the interventions had an insignificant overall impact on tailgating events, with a general significance level of (Wald Chi-Square =1.804, p =0.614) and the smallest pairwise significance value being p=0.258 between phase 3 and phase 4, suggesting a slightly better improvement rate between these phases. Additionally, the difference in the rate of change between high-severity and medium-severity tailgating events was also found to be insignificant, with the interaction term group*phases yielding a significance value of (Wald Chi-Square = 0.033, p =0.998). As shown in the mean profile plots in Figure 4.10, both high and medium-severity tailgating events in the study area were gently impacted by the interventions and were found to be statistically insignificant. These findings are presented in Table 4.8. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.7: Estimates of High and Medium severity tailgating events across phases

Estimates					
			Std.	95% Wald CI	
Group		Mean	Error	Lower	Upper
High_Severity	Phase 1	119.77	34.353	68.27	210.13
	Phase 2	108.35	20.240	75.14	156.26
	Phase 3	116.08	23.947	77.48	173.93
	Phase 4	96.06	15.766	69.64	132.51
Medium_Severity	Phase 1	521.43	86.086	377.28	720.65
	Phase 2	505.96	62.089	397.80	643.53
	Phase 3	532.33	71.813	408.65	693.45
	Phase 4	444.71	51.458	354.47	557.91

Table 4.8: Tests model effects for High and Medium severity tailgating events

Tests of Model Effects			
		Type III	
Source	Wald Chi-Square	df	Sig.
(Intercept)	7311.847	1	0.000
Group	141.289	1	0.000
Phases	1.804	3	0.614
Group * Phases	0.033	3	0.998
Dependent Variable: by_severity_Tailgating Model: (Intercept), Group, Phases, Group * Phases			

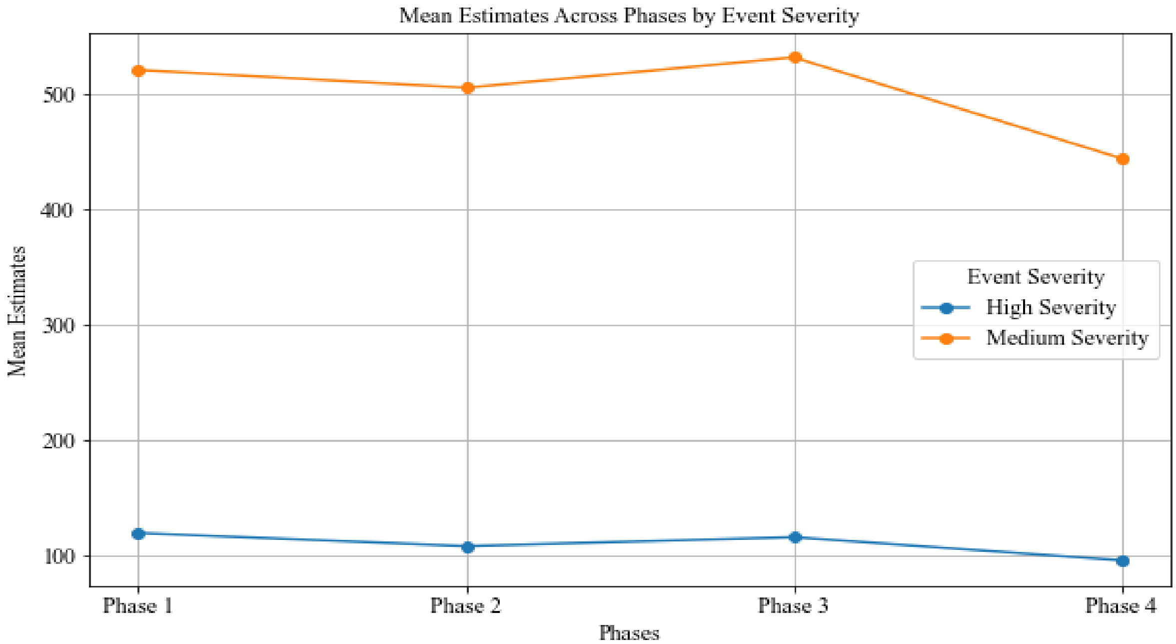

Figure 4.10: Mean profile for High and Medium severity tailgating events across phases

4.4.2.2 Speeding

Similarly, the statistical analysis revealed that the interventions had an insignificant overall impact on speeding events, with a general significance level of (Wald Chi-Square =1.046, p =0.79). The smallest pairwise significance value was p =0.38 between phase 2 and phase 4, indicating a marginally better improvement rate between these phases. Additionally, the difference in the rate of change between high-severity and medium-severity speeding events was also insignificant, with the interaction term group*phases showing a significance value of (Wald Chi-Square = 0.177, p =0.981). As shown in the mean profile plots in Figure 4.11, both high and medium-severity speeding events in the study area were lightly impacted by the interventions and were found to be statistically insignificant. These findings are presented in Table 4.10. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.9: Estimates of High and Medium severity speeding events across phases

Estimates					
			Std.	95% Wald CI	
Group		Mean	Error	Lower	Upper
High_Severity	1	137.69	18.400	105.96	210.13
	2	139.88	16.488	111.03	156.26
	3	134.96	14.796	108.87	173.93
	4	120.69	14.345	95.61	132.51
Medium_Severity	1	57.04	7.236	44.48	720.65
	2	55.84	5.969	45.29	643.53

	3	54.06	5.689	43.98	693.45
	4	52.49	6.274	41.53	557.91

Table 4.10: Tests model effects for High and Medium severity speeding events

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	11428.036	1	0.000
Group	113.567	1	0.000
Phases	1.046	3	0.790
Group * Phases	0.177	3	0.981

Dependent Variable: Speeding_ by Severity
Model: (Intercept), Group, Phases, Group * Phases

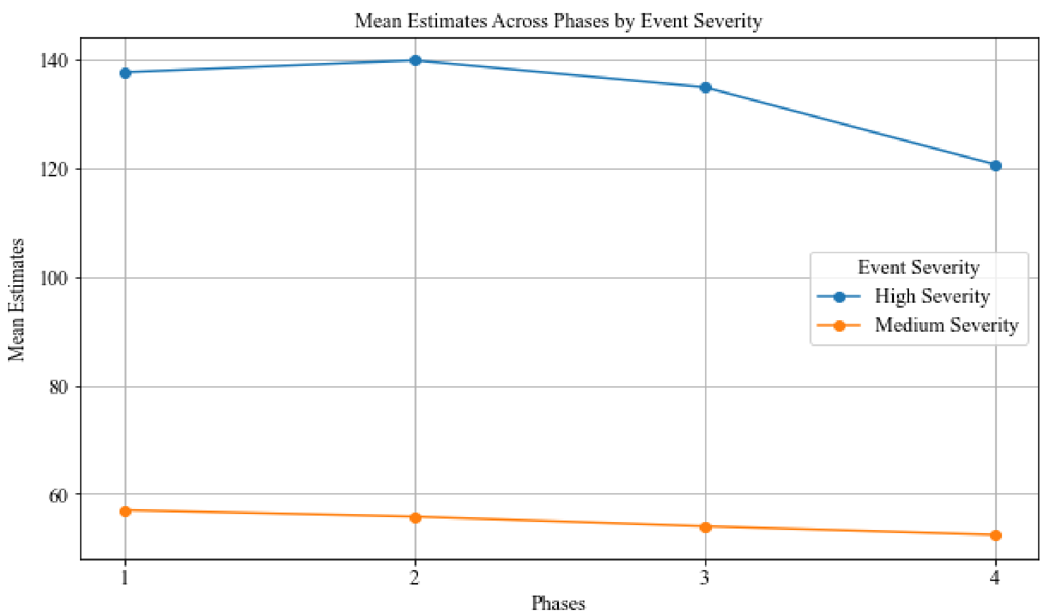

Figure 4.11: Mean profile for High and Medium severity Speeding events across phases

4.4.3 Analysis by Risk Zone

In this section, the data filtered by risk zone from the spatial analysis was utilized to evaluate the effectiveness of the interventions in high and low-risk zones. The primary focus was to determine whether the interventions led to significant differences in the rate of change for tailgating and speeding events across these zones. The analysis identified that the interventions for tailgating events had a more pronounced effect in high-risk zones than in low-risk zones. In contrast, this effect was insignificant for speeding events.

4.4.3.1 Tailgating

Based on the statistical analysis results, the difference in the rate of change between tailgating events in high and low-risk zones was found to be significant, with the interaction term group*phases yielding a significance value of (Wald Chi-Square = 11.342, p=0.01<0.05). Additionally, the interventions significantly impacted tailgating events in high-risk zones, with a significance level of p<0.001 as obtained from the group*phases pairwise comparisons in Table 4.13. In contrast, for tailgating events in the lower-risk zone, the impact of the intervention was insignificant, with the smallest group*phases pairwise significance level of p=0.0923 between phase 1 and phase 4. These findings are presented in Table 4.11. Moreover, as shown in Figure 4.12, the mean of tailgating events in the high-risk zone changed sharply from Phase 1 to Phase 2, and this change continued up to Phase 4, indicating that the events in this zone were impacted by the interventions and were found to be statistically insignificant. Additionally, there were also considerable changes in the low-risk zone. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.11: Tests of model effects for tailgating events in high and low-risk zone

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	7500.540	1	0.000
Group	14.674	1	0.000
Phase	28.062	3	0.000
Group * Phase	11.342	3	0.010

Dependent Variable: Tailgating_by_Risk_zone
Model: (Intercept), Group, Phase, Group * Phase

Table 4.12: Estimates for tailgating events in high and low-risk zone

Estimates					
			Std.	95% Wald CI	
Group	Phases	Mean	Error	Lower	Upper
High Risk_Zone	Phase 1	331.91	52.236	243.81	451.84
	Phase 2	114.53	23.875	76.12	172.33
	Phase 3	111.87	21.441	76.83	162.87
	Phase 4	95.53	18.307	65.62	139.08
Low Risk_Zone	Phase 1	266.92	44.540	192.46	370.19
	Phase 2	218.71	25.767	173.61	275.52
	Phase 3	240.29	40.090	173.27	333.24
	Phase 4	181.39	24.440	139.29	236.21

*Table 4.13: Group*phases interaction pairwise comparisons for tailgating events*

Pairwise Comparisons								
Group*Phase			Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald CI Lower	Upper
High-Risk Zone	1	2	217.38a	57.434	1	0.0002	104.81	329.95
		3	220.04a	56.465	1	0.0001	109.37	330.71
		4	236.38a	55.351	1	0.0000	127.89	344.86
Low-Risk Zone	1	2	48.22	51.456	1	0.3487	-52.64	149.07
		3	26.63	59.925	1	0.6568	-90.82	144.08
		4	85.53	50.805	1	0.0923	-14.05	185.10

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Tailgating by Risk zone
a. The mean difference is significant at the .05 level.

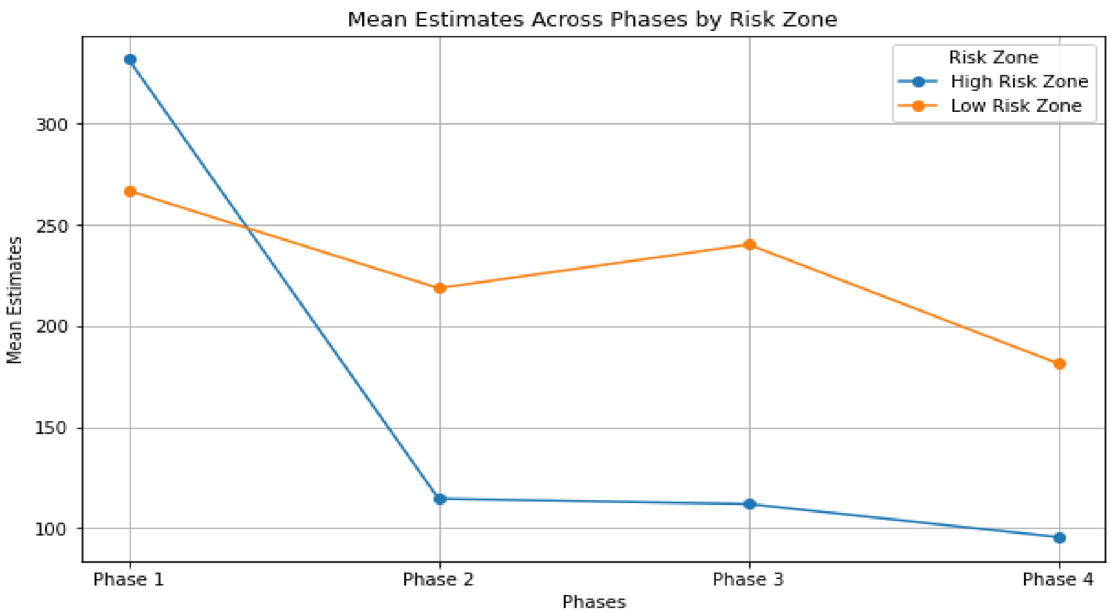

Figure 4.12: Mean Profile for Tailgating events in high and low-risk zones across the phases

4.4.3.2 Speeding

Similarly, the analysis was conducted for speeding events; however, the difference in the rate of change between speeding events in high and low-risk zones was found to be insignificant, with the interaction term group*phases yielding a significance value of (Wald Chi-Square = 0.071, p=0.995), as presented in Table 4.14. Additionally, the impact of the interventions on speeding events in high and low-risk zones was found to be insignificant. As shown in Table 4.16, the smallest pairwise significance levels were p=0.210 and p=0.078 when comparing Phase 1 and Phase 4 in high and low-risk zones, respectively. This indicated that the interventions considerably impacted speeding events in the low-risk zone compared to events in the higher-risk zone, considering a 90% confidence level. The mean profile plots also demonstrated these findings. The high-risk zone defined by the spatial analysis generally covers a smaller area, and the mean count of speeding events was relatively lesser than the low-risk zone as we have no control over the number of drivers and their corresponding number of events in each zone. Further results on the pairwise comparisons between phases and group*phase interaction can be found in Appendix B.

Table 4.14: Tests of model effects for Speeding events in high and low-risk zone

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	4937.554	1	0.000
Group	83.776	1	0.000
Phase	4.430	3	0.219
Group * Phase	0.071	3	0.995
Dependent Variable: Speeding_by_Risk_Zone			
Model: (Intercept), Group, Phase, Group * Phase			

Table 4.15: Estimates for Speeding events in high and low-risk zone

Estimates					
			Std.	95% Wald CI	
Group	Phases	Mean	Error	Lower	Upper
High_Risk	Phase 1	33.93	6.421	23.41	49.16
	Phase 2	28.59	4.890	20.44	39.97
	Phase 3	26.27	4.932	18.18	37.95
	Phase 4	24.02	4.588	16.52	34.93
Low_Risk	Phase 1	90.08	11.139	70.69	114.78
	Phase 2	75.13	8.258	60.57	93.20
	Phase 3	74.21	7.871	60.28	91.36
	Phase 4	65.94	8.001	51.99	83.65

*Table 4.16: Group*phases interaction pairwise comparisons for speeding events*

Pairwise Comparisons								
			Mean				95% Wald CI	
Group*Phase			Difference (I-J)	Std. Error	df	Sig.	Lower	Upper
High-Risk Zone	1	2	5.34	8.071	1	0.508	-10.48	21.16
		3	7.66	8.096	1	0.344	-8.21	23.53
		4	9.90	7.891	1	0.210	-5.56	25.37
Low-Risk Zone	1	2	14.94	13.866	1	0.281	-12.23	42.12
		3	15.87	13.639	1	0.245	-10.87	42.60
		4	24.13	13.714	1	0.078	-2.74	51.01

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Speeding_by_Risk_Zone
a. The mean difference is significant at the .05 level.

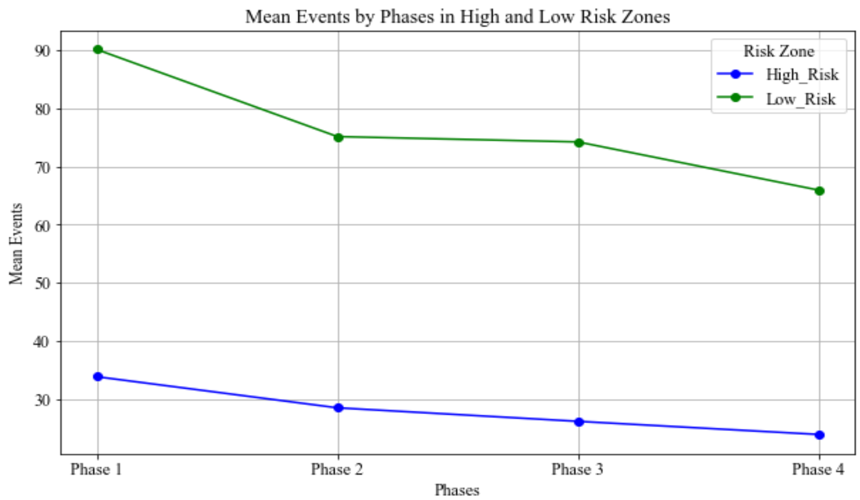

Figure 4.13: Mean profile for speeding events in high and low-risk zones across the phases

4.4.4 Analysis by Road Types

Further analysis was conducted to examine the impact of interventions in high- and low-risk zones classified across different road categories and to determine if the effectiveness of the interventions varied among motorways, primary roads, and secondary roads. The data was relatively sufficient for tailgating events in high- and low-risk zones to be categorized into three road types (motorways, primary roads, and secondary roads). This allowed for a detailed comparison of the intervention's impact across these distinct road types. However, when attempting to categorize speeding events within high-risk zones across the different roadway categories, the data proved insufficient for a robust analysis. Consequently, only tailgating events in high- and low-risk zones were further subdivided into road categories, allowing a thorough evaluation of the intervention's effectiveness across these categories.

The entire speeding dataset was used for speeding events to identify events along motorways, primary roads, and secondary roads. A comparative analysis was then performed to determine if there were any significant differences in the impact of the interventions on speeding events across these road types. This comprehensive approach aimed to provide an understanding of how road type influences the effectiveness of safety interventions, ensuring that the analysis captures the full scope of the intervention's impact across varying traffic environments.

4.4.4.1 Tailgating

Analyzing tailgating events within high and low-risk zones, categorized by roadway types (motorways, primary roads, and secondary roads), provided further insights into the intervention's effectiveness.
For high-risk zones, the Tests of Model Effects in Table 4.17 showed significant overall effects of phases on tailgating events (Wald Chi-Square = 12.033, $p = 0.007<0.05$), indicating that the interventions effectively reduced tailgating events. The pairwise comparisons of the phases in Appendix B also demonstrated several

significant reductions (at 95% confidence level) from Phase 1 to Phases 2, 3, and 4, with mean differences of 23.21 (p = 0.026), 25.61 (p = 0.011), and 29.60 (p = 0.004), respectively. However, the interaction effect of group*phases shown in Table 4.17 was insignificant (Wald Chi-Square = 6.438, p = 0.376), suggesting that the impact of interventions did not differ significantly across the motorway, primary, and secondary roads. Despite this, the pairwise comparisons within the motorway category in Table 4.19 showed a significant reduction between Phase 1 and Phase 4 (mean difference 56.85, p = 0.033<0.05), indicating a meaningful impact of the intervention on motorways. The mean profile plots in Figure 4.14 also demonstrated that the impact of the interventions is more pronounced on motorways than on primary roads and had the least effect on secondary roads. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.17: Tests of model effects for high-risk zone tailgating events in different road categories

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	2615.941	1	0.000
group	93.136	2	0.000
Phases	12.033	3	0.007
group * Phases	6.438	6	0.376

Dependent Variable: High Risk Tailgating_R. Type
Model: (Intercept), group, Phases, group * Phases

Table 4.18: Estimates of high-risk zone tailgating events in different road categories

Estimates					
			Std.	95% Wald CI	
Group	Phases	Mean	Error	Lower	Upper
MW	Phase 1	80.65	25.647	43.24	150.41
	Phase 2	25.38	6.787	15.03	42.87
	Phase 3	24.75	5.570	15.92	38.47
	Phase 4	23.80	7.163	13.19	42.93
PR	Phase 1	103.33	24.139	65.37	163.34
	Phase 2	94.44	19.013	63.64	140.12
	Phase 3	77.84	16.205	51.76	117.06
	Phase 4	81.89	20.585	50.03	134.03
SR	Phase 1	25.04	6.079	15.56	40.30
	Phase 2	19.63	4.717	12.26	31.44
	Phase 3	19.87	4.613	12.61	31.32
	Phase 4	13.47	2.859	8.88	20.42

*Table 4.19: Group*phases pairwise comparison of high-risk zone tailgating events in different road categories*

Pairwise Comparisons

Group*Phases			Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald CI Lower	Upper
MW	1	2	55.27[a]	26.530	1	0.037	3.27	107.27
		3	55.90[a]	26.245	1	0.033	4.46	107.34
		4	56.85[a]	26.629	1	0.033	4.66	109.04
PR	1	2	8.90	30.728	1	0.772	-51.33	69.12
		3	25.50	29.074	1	0.381	-31.49	82.48
		4	21.44	31.724	1	0.499	-40.73	83.62
SR	1	2	5.41	7.695	1	0.482	-9.67	20.49
		3	5.17	7.631	1	0.498	-9.79	20.13
		4	11.57	6.718	1	0.085	-1.59	24.74

Pairwise comparisons of estimated marginal means are based on the original scale of the dependent variable, High-Risk Tailgating_R. Type
a. The mean difference is significant at the .05 level.

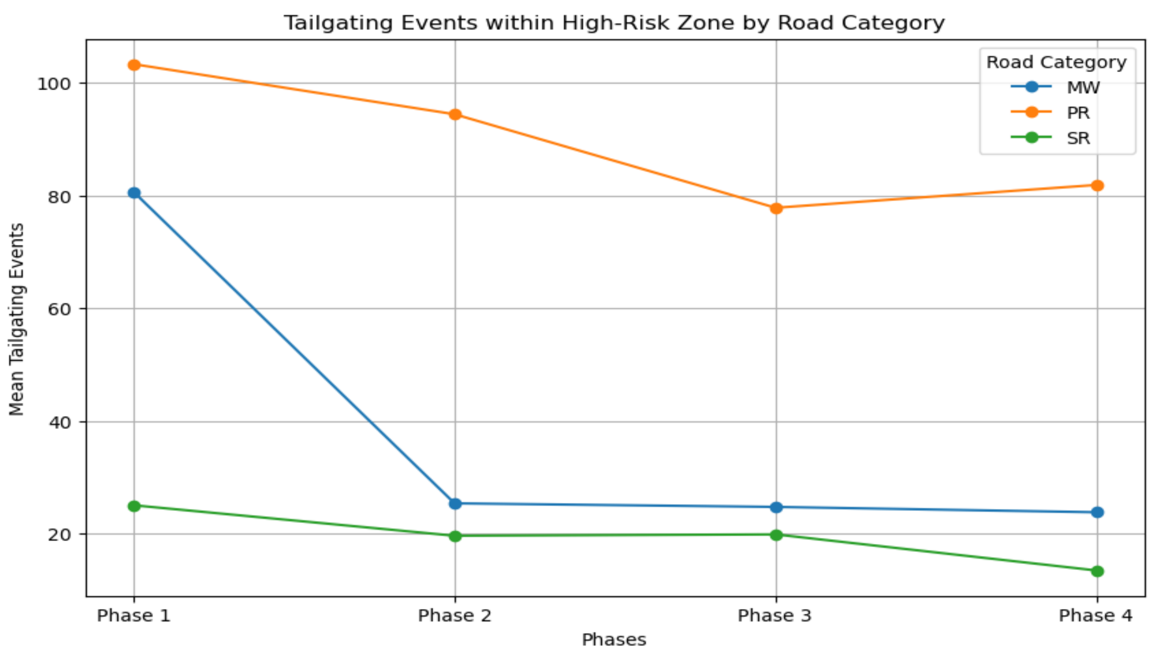

Figure 4.14: Mean profile for high-risk zone tailgating events in different road categories

On the other hand, in low-risk zones, the overall effect of phases on tailgating events was insignificant (Wald Chi-Square = 3.637, p = 0.303), as shown in Table 4.20, indicating that the interventions did not significantly reduce tailgating events. The interaction effect of group*phases was also insignificant (Wald Chi-Square = 1.636, p = 0.950), showing no significant differences in the intervention's impact across different road types. Moreover, pairwise comparisons within low-risk zones demonstrated that the phases did not significantly affect tailgating events in any specific group (motorway, primary, and secondary roads), even though there was a visible reduction in tailgating events along primary and secondary roads. These findings are presented in Table 4.22. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.20: Tests of model effects for low-risk zone tailgating events in different road categories

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	4414.087	1	0.000
group	38.550	2	0.000
phase	3.637	3	0.303
group * phase	1.636	6	0.950
Dependent Variable: Low-Risk Tailgating_R. Type			
Model: (Intercept), group, phase, group * phase			

Table 4.21: Estimates of low-risk zone tailgating events in different road categories

Estimates					
			Std.	95% Wald CI	
Group	Phases	Mean	Error	Lower	Upper
MW	Phase 1	85.68	24.843	48.53	151.24
	Phase 2	81.60	14.938	57.00	116.82
	Phase 3	120.10	44.965	57.66	250.16
	Phase 4	73.08	18.741	44.21	120.80
PR	Phase 1	75.22	14.661	51.34	110.22
	Phase 2	65.86	11.458	46.83	92.62
	Phase 3	58.96	10.057	42.21	82.37
	Phase 4	53.28	10.091	36.75	77.22
SR	Phase 1	43.36	6.208	32.75	57.40
	Phase 2	40.65	6.134	30.24	54.64
	Phase 3	36.58	5.288	27.56	48.56
	Phase 4	29.74	4.422	22.23	39.80

Table 4.22: Tests of model effects for low-risk zone tailgating events in different road categories

Pairwise Comparisons								
			Mean Difference	Std.			95% Wald CI	
Group*phase			(I-J)	Error	df	Sig.	Lower	Upper
MW	1	2	4.08	28.989	1	0.888	-52.74	60.89
		3	-34.42	51.371	1	0.503	-135.11	66.26
		4	12.60	31.119	1	0.686	-48.39	73.59
PR	1	2	9.37	18.608	1	0.615	-27.10	45.84
		3	16.26	17.779	1	0.360	-18.58	51.11

		4	21.95	17.798	1	0.217	-12.93	56.83
SR	1	2	2.71	8.727	1	0.756	-14.39	19.82
		3	6.78	8.154	1	0.406	-9.21	22.76
		4	13.62	7.621	1	0.074	-1.32	28.55

Pairwise comparisons of estimated marginal means based on the original scale of the dependent variable, Low-Risk Tailgating_R. Type
a. The mean difference is significant at the .05 level.

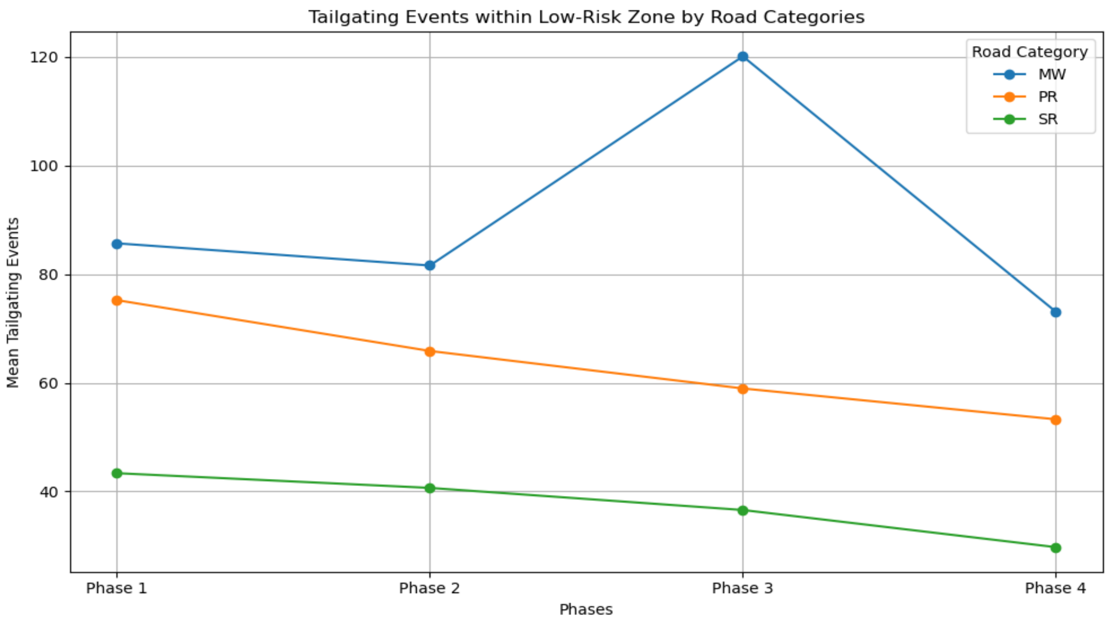

Figure 4.15: Mean profile for low-risk zone tailgating events in different road categories

4.4.4.2 Speeding

The analysis by road types was extended to speeding events by filtering the entire speeding dataset along the three road categories (motorways, primary roads, and secondary roads) to investigate further the effect of the phases on speeding events. The test of the model effects in Table 4.23 demonstrated that the phases had no significant effect on speeding events overall (Wald Chi-Square = 1.752, p = 0.625), indicating that the interventions did not significantly reduce speeding events. Additionally, the interaction effect of group*phases was insignificant (Wald Chi-Square = 0.735, p = 0.994), showing no significant differences in the intervention's effectiveness across different road types. Pairwise comparisons in Table 4.25 also demonstrated that the phases did not significantly affect speeding events across motorways, primary roads, and secondary roads. The mean profile plots in Figure 4.16 also confirmed that there was only a minor change in the means of the intervention phases compared to the baseline phase, with a slightly better change in the primary roads category. Further details on the significance of pairwise comparisons between phases and group*phases interactions can be found in Appendix B.

Table 4.23: Tests of model effects for speeding events in different road categories

Tests of Model Effects			
	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	4072.129	1	0.000
Group	80.076	2	0.000
Phases	1.752	3	0.625
Group * Phases	0.735	6	0.994

Dependent Variable: Speeding_by_road_types
Model: (Intercept), Group, Phases, Group * Phases

Table 4.24: Tests of model effects for speeding events in different road categories

Estimates					
			Std.	95% Wald CI	
Group	Phases	Mean	Error	Lower	Upper
MW	Phase 1	16.72	4.788	9.54	29.31
	Phase 2	14.13	2.276	10.30	19.37
	Phase 3	16.18	3.053	11.18	23.42
	Phase 4	13.82	3.496	8.42	22.69
PR	Phase 1	59.14	11.576	40.30	86.80
	Phase 2	55.08	9.882	38.75	78.29
	Phase 3	54.46	8.174	40.58	73.09
	Phase 4	49.32	7.892	36.04	67.49
SR	Phase 1	34.22	5.487	25.00	46.86
	Phase 2	35.59	5.088	26.90	47.10
	Phase 3	36.53	5.419	27.31	48.86
	Phase 4	28.45	3.670	22.09	36.63

*Table 4.25: Group*phases pairwise comparisons for speeding events in different road categories*

Pairwise Comparisons								
			Mean				95% Wald CI	
Group*Phases			Difference (I-J)	Std. Error	df	Sig.	Lower	Upper
MW	1	2	2.59	5.302	1	0.625	-7.80	12.98
		3	0.54	5.679	1	0.924	-10.59	11.67
		4	2.90	5.929	1	0.625	-8.72	14.52
PR	1	2	4.06	15.221	1	0.790	-25.77	33.89
		3	4.68	14.171	1	0.741	-23.10	32.46
		4	9.82	14.010	1	0.483	-17.64	37.28

SR	1	2	-1.37	7.483	1	0.855	-16.03	13.30
		3	-2.31	7.711	1	0.765	-17.42	12.81
		4	5.78	6.601	1	0.382	-7.16	18.71

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Speeding_by_road_types
a. The mean difference is significant at the .05 level.

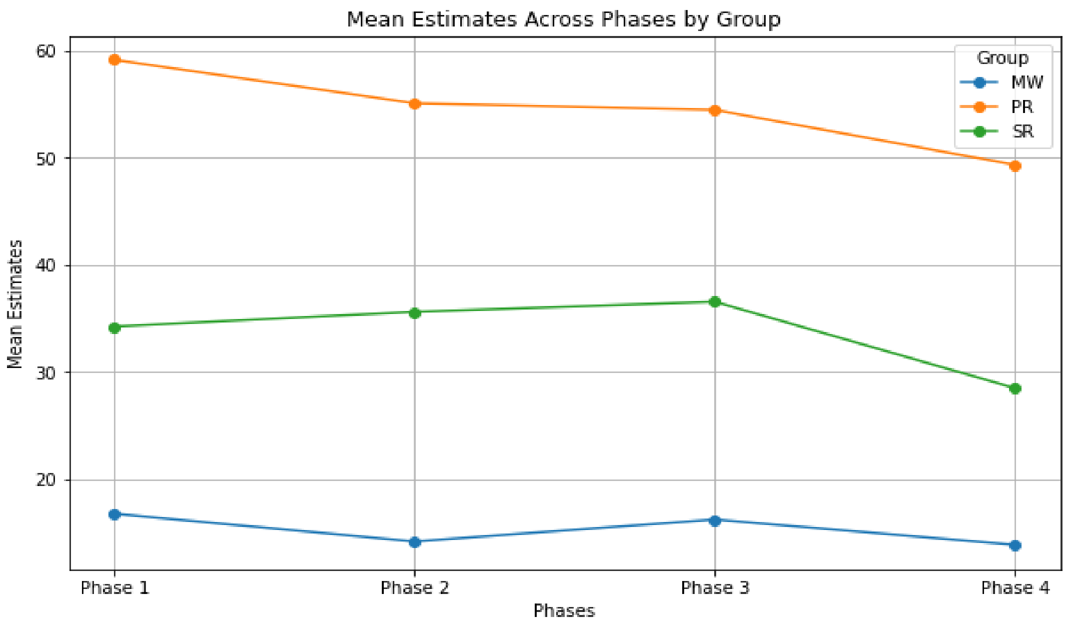

Figure 4.16: Mean profile for Speeding events in different road categories

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The study aimed to evaluate the effectiveness of the i-DREAMS intervention in reducing risky driving behaviors, specifically tailgating and speeding, across different risk zones and roadway categories. Through spatial and statistical analyses, the study addressed the research questions and objectives designed to assess the intervention's impact comprehensively.

Spatial Analysis and Distribution of Risky Driving Events

The spatial analysis identified high-risk zones and hotspots for tailgating and speeding events, primarily concentrated in major cities and high-traffic areas in the Limburg province of Belgium. Specifically, the hotspots for these events were centered around Hasselt and Genk, predominantly occurring on motorways and primary roads where more traffic interactions exist. This detailed spatial mapping provided a crucial foundation for subsequent statistical analyses, enabling a focused investigation of the intervention's impact across different contexts.

Effectiveness of Intervention Phases

- Overall, the intervention phases reduced the number of tailgating and speeding events from the baseline phase in several instances, as detailed in the discussion section of this study. However, based on the tests conducted, only some of these reductions reached statistical significance, specifically in the case of tailgating events in high-risk zones. This indicated that the impact of the intervention phases varies based on risky driving events (performance objectives) and the risk zone.

- Further statistical analysis based on event severity level was conducted to address the question of whether there was a difference in the effect of the intervention phases in impacting high and medium-severity tailgating and speeding events. The results demonstrated that the difference in the rate of change between high-severity and medium-severity tailgating and speeding events was found to be insignificant, indicating that there was no significant difference in the impact of the intervention phases based on the severity of events.

Comparing High and Low-Risk Zones

- Another test was conducted to determine the difference in the intervention impacts in the different risk zones of the study area. The statistical analysis revealed that the intervention was significantly more effective in high-risk zones for tailgating events, as evidenced by the significant interaction term for group*phases (p =0.01). The Tests of Model Effects demonstrated that phases 2, 3, and 4 significantly reduced tailgating events compared to phase 1, with the most substantial reductions observed on motorways. In contrast, the impact in low-risk zones was not significant. On the other hand, neither high-risk nor low-risk zones showed significant improvements in speeding events. This inconsistency in the effectiveness of the interventions can be attributed to different factors, one of which can be the tendency of the participants to engage in speeding behavior, as depicted by the higher proportion of high-severity speeding events in the data. In addition, this can also be related to the thresholds defined for the speeding and tailgating events.

Effectiveness Across Roadway Categories:

- Another analysis was made to answer the question of whether the intervention effectiveness varies across the different road categories. The results of tailgating events in high-risk zones showed significant reductions across the motorway, primary, and secondary roads, particularly notable on motorways. However, the interaction effects across different road types were insignificant for both tailgating and speeding events, indicating no variation in intervention effectiveness based on road category.

Improving Intervention Measures:

The findings suggested that while the i-DREAMS intervention effectively reduced tailgating events in high-risk zones, it had a limited statistically significant impact on speeding events. These results suggest a need to re-evaluate the intervention strategies, particularly considering the thresholds for speeding events and developing targeted approaches to address high-severity speeding behaviors, which could potentially improve the overall effectiveness of the intervention.

5.2 Recommendations

The spatial and quantitative analysis of this study found that risky driving behaviors persisted even with the presence of the interventions, indicating the persistent nature of risky driving behavior among the participants. This finding aligns with previous research, which indicated that more than half of the Belgian drivers in the i-DREAMS field trials had recent traffic offenses, predominantly for speeding (Brown L. et al., 2023). These results underline the necessity for stronger policies against aggressive driving behaviors.

In line with this study, the following practical recommendations are drawn: -

Lowering Thresholds for In-vehicle Warnings

- Adjust the speed thresholds at which the i-DREAMS system triggers in-vehicle warnings. Lowering the threshold may prompt earlier interventions before speeding or tailgating becomes severe allowing drivers to have more time to adjust their behavior.

Strengthening the Monitoring Aspects of the Intervention

- Precise data collection that considers factors such as weather conditions, traffic density, and road type and geometry when collecting speed limit data and determining thresholds for warnings, which can increase trust and acceptability of the technology.

Increased Use of Gamification: Strengthen the gamification aspect of the post-trip feedback through positive reinforcement and by introducing more engaging challenges and rewards focused specifically on adherence to speed limits.

Targeted Interventions and Enhanced Enforcement in High-Risk Areas

- Implement targeted interventions such as road signage improvements (variable speed limit signage) that provide clear and real-time speed limit information. Additionally, increased enforcement measures, such as speed cameras and radar, should be implemented, particularly in areas identified as high-risk zones for speeding, particularly around major cities and along motorways.

Strategies for Specific Driving Behaviors

- Based on the findings in this study, it is necessary to develop tailored strategies focusing on specific driving behaviors. The study findings also suggest i-DREAMS interventions (in-vehicle warning, post-trip smartphone-based feedback, and post-trip web-based gamified interventions) targeting speeding behaviors to be reassessed and potentially redesigned to address their unique challenges (high severity) more effectively.
- Establish a continuous monitoring system to identify trends and patterns in speeding behavior using the data collected from the i-DREAMS technology and adjust speed limits accordingly in high-risk areas. This data can then be used to iteratively refine and adapt the interventions, ensuring they remain effective in changing driving behaviors over time.

5.3 Limitations of the Study and Future Research

This study is confined to certain limitations in scope and methodology, which should be considered when interpreting the findings:

- Geographical Scope: The study is based on data from car drivers within Belgium as part of the i-DREAMS project. Although the i-DREAMS initiative spans five European countries and multiple transport modes, this research focuses solely on Belgium, limiting the generalizability of the results to other regions and contexts.

- Focus on Specific Risky Events: The study exclusively evaluates the impact of interventions on speeding and tailgating events. While these are significant aspects of risky driving behavior, excluding other potential risky driving events limits the comprehensiveness of the findings.

- Socio-Demographic Factors: This research did not account for the correlations between socio-demographic factors and risky driving events, despite previous findings within the i-DREAMS project suggesting no significant correlation between such factors and speeding or tailgating. Accounting for these variables, how the interventions vary based on these factors.

- Traffic Volume Assumptions: The study assumes a constant traffic volume across different phases, neglecting potential variations in traffic flow that could influence the results. This assumption may lead to an incomplete understanding of the intervention's impact, as changes in traffic volume can significantly affect driving behavior and safety outcomes.

To address the limitations identified and enhance the robustness of future studies, the following areas are recommended for further research:

- Expanded Geographical Coverage: Future studies should extend the analysis to include all five European countries involved in the i-DREAMS project (UK, Germany, Portugal, Greece, and Belgium). This broader scope would enable more generalizable results and facilitate geographic comparisons, helping to identify contexts where the interventions are most effective.

- Inclusion of Additional Risky Driving Events: Research should be expanded to consider a wider range of risky driving behaviors beyond speeding and tailgating. Future studies can provide a more holistic approach to improving road safety by encompassing a broader spectrum of events.

- Consideration of Traffic Flow Variations: Future research should take into account variations in traffic flow across different phases of intervention. By incorporating traffic data, researchers can better understand the dynamic effects of traffic volume on driving behavior and the effectiveness of safety interventions.

- Detail Analysis of Road Infrastructure Parameters: Road infrastructure elements such as lane width, number of lanes, curves, and other geometric characteristics significantly influence driving behavior. Future studies should integrate these variables to understand how road design affects risky driving behaviors comprehensively.

- Behavioral Studies and Alternative Approaches: Complementing spatial and statistical analyses with behavioral studies will provide deeper insights into the psychological and sociological factors underlying aggressive driving behaviors, such as speeding and tailgating. This can be done by incorporating question-based analysis and incorporating those factors.

REFERENCES

Abd Rahman, N. I., Dawal, S. Z. M., & Yusoff, N. (2020). Driving mental workload and performance of aging drivers. Transportation Research Part F: Traffic Psychology and Behaviour, 69, 265–285.

Adanu, E. K., Smith, R., Powell, L., & Jones, S. (2017). Multilevel analysis of the role of human factors in regional disparities in crash outcomes. Accident Analysis & Prevention, pp. 109, 10–17. https://doi.org/10.1016/J.AAP.2017.09.022

Adell, E., Várhelyi, A., & Fontana, M. dalla. (2011). The effects of a driver assistance system for safe speed and safe distance - A real-life field study. Transportation Research Part C: Emerging Technologies, 19(1), 145–155. https://doi.org/10.1016/j.trc.2010.04.006

Adminaité-Fodor, D., Stipdonk, H., Thomas, P., Yannis, G., Avenoso, A., Jost, G., Carson, J., & ETSC. (2021). Ranking EU progress on road safety. June. www.etsc.eu/pin

Aidman, E., Chadunow, C., Johnson, K., & Reece, J. (2015). Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. Accident Analysis and Prevention, 81(August), 8–13.

Ahn, Y., & Kang, K. (2019). Identification of high-risk zones in road network using smartphone-based driving behavior data. Accident Analysis & Prevention, 126, 100-108.

Alam, M. S., & Tabassum, N. J. (2023). Spatial pattern identification and crash severity analysis of road traffic crash hot spots in Ohio. Heliyon, 9(5), e16303. https://doi.org/10.1016/j.heliyon.2023.e16303

Alemu, D. (2024). Vehicle control behavior of drivers as a proxy indicator for crashes (Master's thesis, Hasselt University). School of Transportation Science.

Alkaabi, K. (2023). Identifying hotspot areas for traffic accidents and analyzing drivers' behaviors and road accidents. Transportation Research Interdisciplinary Perspectives, 22(August), 100929. https://doi.org/10.1016/j.trip.2023.100929

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

Bagdadi, O. (2013). Assessing safety critical braking events in naturalistic driving studies. Transportation Research Part F: Traffic Psychology and Behaviour, 16, 117–126. https://doi.org/10.1016/j.trf.2012.08.006

Brijs, K., Adnan, M., Ross, V., Cuenen, A., Vanrompay, Y., Khattak, MW, ... & Brijs, T. (2023). Effectiveness of real-time and post-trip interventions from the H2020 i-DREAMS naturalistic driving project: A Sneak Preview. *Transportation Research Procedia* , *72* , 2133-2140.

Brijs, K. (2020). D3.3. Toolbox of recommended interventions to assist drivers in maintaining a safety tolerance zone Project identification Document summary. 814761, 0–181.

Brown, L., Talbot., R., Khattak, M. W., Petraki, V., Gruden, C. et al. (2023). Effectiveness Evaluation of the Interventions. Deliverable 7.2 of the EC H2020 project i-DREAMS.

Cai, M., Yazdi, M. A. A., Mehdizadeh, A., Hu, Q., Vinel, A., Davis, K., Xian, H., Megahed, F. M., & Rigdon, S. E. (2021). The association between crashes and safety-critical events: Synthesized evidence from crash reports and naturalistic driving data among commercial truck drivers. Transportation Research Part C: Emerging Technologies, 126(October 2019), 103016. https://doi.org/10.1016/j.trc.2021.103016

Cameron, A. C., & Trivedi, P. K. (2013). Regression Analysis of Count Data. Cambridge University Press.

Charly, A., & Mathew, T. V. (2023). Identifying risky driving behavior: a field study using instrumented vehicles. Transportation Letters, 00(00), 1–15. https://doi.org/10.1080/19427867.2023.2233782

Daniels, S. (2022). Road safety strategy in a regionalized country.

Dijksterhuis, C., Lewis-Evans, B., Jelijs, B., De Waard, D., Brookhuis, K., & Tucha, O. (2015). The impact of immediate or delayed feedback on driving behavior in a simulated Pay-As-You-Drive system. Accident Analysis and Prevention, 75, 93–104. https://doi.org/10.1016/j.aap.2014.11.017

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences, 113(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113.

Dobson, A. J., & Barnett, A. G. (2018). An Introduction to Generalized Linear Models (4th ed.). CRC Press.

Dumbaugh, E., & Rae, R. (2009). Safe urban form: Revisiting the relationship between community design and traffic safety. Journal of the American Planning Association, 75(3), 309–339.

Dupont, E., Papadimitriou, E., Martensen, H., & Yannis, G. (2013). Multilevel analysis in road safety research. Accident Analysis and Prevention, pp. 60, 402–411. https://doi.org/10.1016/j.aap.2013.04.035

Elvik, R. (2013). A meta-analysis of the effects of road design improvements on accident risk. Accident Analysis & Prevention, pp. 50, 35–47.

European Commission. (2020). Next steps towards Vision Zero.

Field, A. (2009). Discovering Statistics Using SPSS (3rd ed.). Sage Publications.

Folla, K., & Kaselouris, K. (2023). Country Profile Belgium. Road Safety Observatory. Brussels, European Commission, Directorate General for Transport.

Freed, S. A., Ross, L. A., Gamaldo, A. A., & Stavrinos, D. (2021). Use multilevel modeling to examine the variability of distracted driving behavior in naturalistic driving studies. Accident Analysis and Prevention, 152. https://doi.org/10.1016/j.aap.2021.105986

Gitlin, L. N., & Czaja, S. J. (2015). Behavioral Intervention Research: Designing, Evaluating, and Implementing. Behavioral Intervention Research: Designing, Evaluating, and Implementing, 2015, 5–7. https://doi.org/10.1891/9780826126597

Gomes, M. J. T. L., Cunto, F., & da Silva, A. R. (2017). Geographically weighted negative binomial regression applied to zonal level safety performance models. *Accident Analysis & Prevention*, *106*, 254-261.

Guo, F., & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. Accident Analysis & Prevention, 61, 3-9.

Guo, F. (2019). Statistical Methods for Naturalistic Driving Studies. Annual Review of Statistics and Its Application. https://doi.org/10.1146/annurev-statistics

Guo, F., Klauer, S. G., Hankey, J. M., & Dingus, T. A. (2010). Near crashes as crash surrogate for naturalistic Driving Studies. Transportation Research Record, 2147, 66–74. https://doi.org/10.3141/2147-09

Haddon, W. (1968). The changing approach to the epidemiology. Prevention, and amelioration of trauma 58(8), 1431–1438.

Hancox, G. (2020). D5.1. Organization of the driving simulator and on-road trial experiments in i-DREAMS. 814761, 1–84.

Hancox, G. et al. (2021). Description of the on-road driving trials for identifying safety tolerance zones and the performance of in-vehicle interventions. Deliverable 5.3 of the EC H2020 project i-DREAMS.

Hilbe, J. M. (2011). Negative Binomial Regression. Cambridge University Press.

i-DREAMS. (2019, May 21). i-DREAMS. Retrieved from Objectives. https://idreamsproject.eu/wp/project/

Jayan, K., & Muruganantham, B. (2021). Advanced Driver Assistance System technologies and its challenges toward the development of autonomous vehicle. In *Intelligent Computing and Applications: Proceedings of ICICA 2019* (pp. 55-72). Springer Singapore.

Katrakazas, C., Al Haddad, C., Taveria, R., & Kaiser, S., Mateus, T. (2020a) Framework for operational design of experimental work in i-DREAMS. Deliverable 3.1 of the EC H2020 project i-DREAMS.

Katrakazas, C., Michelaraki, E., & Yannis, G. (2020b). D3.2. Toolbox of recommended data collection tools and monitoring methods and a conceptual definition of the Safety Tolerance Zone. 814761.

Katrakazas, C., Michelaraki, E., Yannis, G., Cuenen, A., Brijs, K., Brijs, T., Fitness, A., Talbot, R., Hancox, G., & Gruden, C. (2020c). Methodology for the evaluation of interventions (Deliverable No. 7.1). NTUA. https://www.project-idreams.eu/.

Keall, M. D., Frith, W. J., & Patterson, T. L. (2004). The influence of roadside advertising signs on driver eye movements and driving performance. Accident Analysis & Prevention, 36(3), 491-497.

Ke, C., He, S., & Qin, Y. (2023). Comparison of natural breaks method and frequency ratio dividing attribute intervals for landslide susceptibility mapping. Bulletin of Engineering Geology and the Environment, 82(384). https://doi.org/10.1007/s10064-023-03392-0

Khattak, M. W., Brijs, K., Tran, T. M., Trinh, T. A., Vu, A. T., & Brijs, T. (2024). Acceptance towards advanced driver assistance systems (ADAS): a validation of the unified model of driver acceptance (UMDA) using structural equation modelling. Transportation research part F: traffic psychology and behaviour, 105, 284-305.

Khattak, M. W., De Backer, H., De Winne, P., Brijs, T., & Pirdavani, A. (2024). Analysis of road infrastructure and traffic factors influencing crash frequency: Insights from generalized Poisson models. Infrastructures, 9(47). https://doi.org/10.3390/infrastructures9030047

Kircher, K., Ahlstrom, C., & Kircher, A. (2007). Effects of driver feedback signs on speeding behavior: a review. Transportation Research Part F: Traffic Psychology and Behavior, 10(6), 459-472.

Kononov, J., Bailey, B., & Allery, B. K. (2008). Relationships between safety, traffic density, and number of lanes on urban freeways. Transportation Research Record: Journal of the Transportation Research Board, 2083(1), 26-39.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Human Factors, 44(2), 314–334. https://doi.org/10.1518/0018720024497844

Leslie, A. J., Kiefer, R. J., Meitzner, M. R., & Flannagan, C. A. (2021). Field effectiveness of general motors advanced driver assistance and headlighting systems. *Accident Analysis & Prevention*, *159*, 106275.

Li, Y., Tian, Z., Zhou, Y., & Li, Y. (2019). The effects of in-vehicle driver behavior feedback systems on risky driving behaviors: A systematic review and meta-analysis. Accident Analysis & Prevention, 124, 106-119.

Mohammed, F. A. (2022). Analyzing the effects of interventions and socio-demographic factors on driving risk events in the i-DREAMS trial experiments (Master's thesis, Hasselt University). School of Transportation Science.

Mohammadi, M. A., Samaranayake, V. A., & Bham, G. H. (2014). Crash frequency modeling using negative binomial models: An application of generalized estimating equation to longitudinal data. Analytic Methods in Accident Research, 2, 52-69.

Naji, H. A. H., Xue, Q., Lyu, N., Wu, C., & Zheng, K. (2018). Evaluating the driving risk of near-crash events using a mixed-ordered logit model. Sustainability (Switzerland), 10(8). https://doi.org/10.3390/su10082868

Nyimbili, P. H., & Erden, T. (2020). A hybrid approach integrating entropy-AHP and GIS for suitability assessment of urban emergency facilities. *ISPRS International Journal of Geo-Information*, *9*(7), 419.

Olleja, P., Bärgman, J., & Lubbe, N. (2022). Can non-crash naturalistic driving data be an alternative to crash data for use in virtual assessment of the safety performance of automated emergency braking systems? Journal of Safety Research, 83, 139-151. https://doi.org/10.1016/j.jsr.2022.08.011

Ouimet, M. C., Brown, T. G., Guo, F., Klauer, S. G., Simons-Morton, B. G., Fang, Y., Lee, S. E., Gianoulakis, C., & Dingus, T. A. (2014). Higher crash and near-crash rates in teenaged drivers with lower cortisol response an 18-month longitudinal, naturalistic study. JAMA Pediatrics, 168(6), 517–522. https://doi.org/10.1001/jamapediatrics.2013.5387

Pande, A., Chand, S., Saxena, N., Dixit, V., Loy, J., Wolshon, B., & Kent, J. D. (2017). A preliminary investigation of the relationships between historical crash and naturalistic driving. Accident Analysis and Prevention, 101, 107–116. https://doi.org/10.1016/j.aap.2017.01.023

Papadimitriou, E., Argyropoulou, A., Tselentis, D. I., & Yannis, G. (2019). Analysis of driver behavior through smartphone data: The case of mobile phone use while driving. Safety Science, 119, 91–97.

Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. Journal of Korean Academy of Nursing, 43(2), 154-164. https://doi.org/10.4040/jkan.2013.43.2.154

Pembuain, A., Priyanto, S., & Suparma, L. B. (2019). The Effect of Road Infrastructure on Traffic Accidents. Advances in Engineering Research, 186, 147-153.

Rahman, M. M., Teo, S. H., & Gan, L. M. (2016). Speeding driving behavior: Age and gender experimental analysis. MATEC Web of Conferences, 74, 5–8. https://doi.org/10.1051/matecconf/20167400030

Regev, S., Rolison, J. J., & Moutari, S. (2018). Crash risk by driver age, gender, and time of day using a new exposure methodology. Journal of Safety Research, 66, 131-140. https://doi.org/10.1016/j.jsr.2018.07.002

Reimer, B., Mehler, B., Wang, Y., Coughlin, J. F., & D'Ambrosio, L. A. (2012). Impact of age, experience, and training on driving performance. Traffic Injury Prevention, 13(sup1), 69-76.

Rhee, K. A., Kim, J. K., Lee, Y. I., & Ulfarsson, G. F. (2016). Spatial regression analysis of traffic crashes in Seoul. *Accident Analysis & Prevention*, *91*, 190-199.

Roberts, S. C., Ghazizadeh, M., & Lee, J. D. (2012). Warn me now or inform me later: Drivers' acceptance of real-time and post-drive distraction mitigation systems. International Journal of Human-Computer Studies, 70(12), 967–979.

Schagen, V., Agberg, I., & Fridulv. (2012). The Potential Benefits of Naturalistic Driving for Road Safety Research: Theoretical and Empirical Considerations and Challenges for the Future. Procedia - Social and Behavioral Sciences, 48, 692–701.

Shaon, M. R. R. (2019). Identifying and incorporating driver behavior variables into crash prediction models. Dissertation Abstracts International: Section B: The Sciences and Engineering, 80(11).

Srinivasan, R., Gross, F., & Bahar, G. (2016). Reliability of Safety Management Methods: Safety Effectiveness Evaluation (FHWA-SA-16-040). Federal Highway Administration. https://safety.fhwa.dot.gov

Talbot, R., Hancox, G., Pilkington, F. C., Filtness, A., Brijs, K., Brijs, T., Ross, V., Katrakazas, C., Haddad, C. al, Taveria, R., Kaiser, S., & Mateus, T. (2020). Framework for operational design of experimental work in i -DREAMS Project identification Document summary. 814761, 1–90.

Tarko, A. P., Saunier, N., & Sayed, T. (2009). Surrogate measures of safety ANB20 (3) Subcommittee on Surrogate Measures of Safety ANB20 Committee on Safety Data Evaluation and Analysis.

Vagenas, D., & Totsika, V. (2018). Modeling correlated data: Multilevel models and generalized estimating equations and their use with data from research in developmental disabilities. Research in Developmental Disabilities, 81, 1-11. https://doi.org/10.1016/j.ridd.2018.04.010

Vanrompay, Y., Donders, E., Fortsakis, P., Brijs, T., Brijs, K., & Wets, G. (2020). A smartphone app (Android) for personalized driving behavioural feedback. Deliverable 4.5 of the EC H2020 project i-DREAMS.

World Health Organization. (2018). Global Status Report on Road Safety (Vol. 11, Issue 1).

Xu, P., & Huang, H. (2015). Modeling crash spatial heterogeneity: Random parameter versus geographic weighting. *Accident Analysis & Prevention*, *75*, 16-25.

Xu, Y., Bao, S., & Pradhan, A. K. (2021). Modeling drivers' reaction when being tailgated: A Random Forests Method. Journal of Safety Research, 78, 28–35. https://doi.org/10.1016/j.jsr.2021.05.004

Yan, X., Wu, C., Shi, J., & Zhang, H. (2017). Evaluation of the effectiveness of driver assistance systems in reducing lane departure crashes. Journal of Safety Research, 63, 89-98.

Yang, D., Xie, K., Ozbay, K., & Yang, H. (2021). Fusing crash data and surrogate safety measures for safety assessment: Development of a structural equation model with conditional autoregressive spatial effect and random parameters. *Accident Analysis & Prevention*, *152*, 105971.

Zandi, K., Tavakoli Kashani, A., & Okabe, A. (2023). Influence of traffic parameters on the spatial distribution of crashes on a freeway to increase safety. Sustainability, 15(1), 493. https://doi.org/10.3390/su15010493

Zhao, G., & Wu, C. (2013). Effectiveness and acceptance of the intelligent speeding prediction system (ISPS). Accident Analysis and Prevention, 52, 19–28

APPENDIX A

Q-GiS spatial Analysis Outputs

1. Road networks

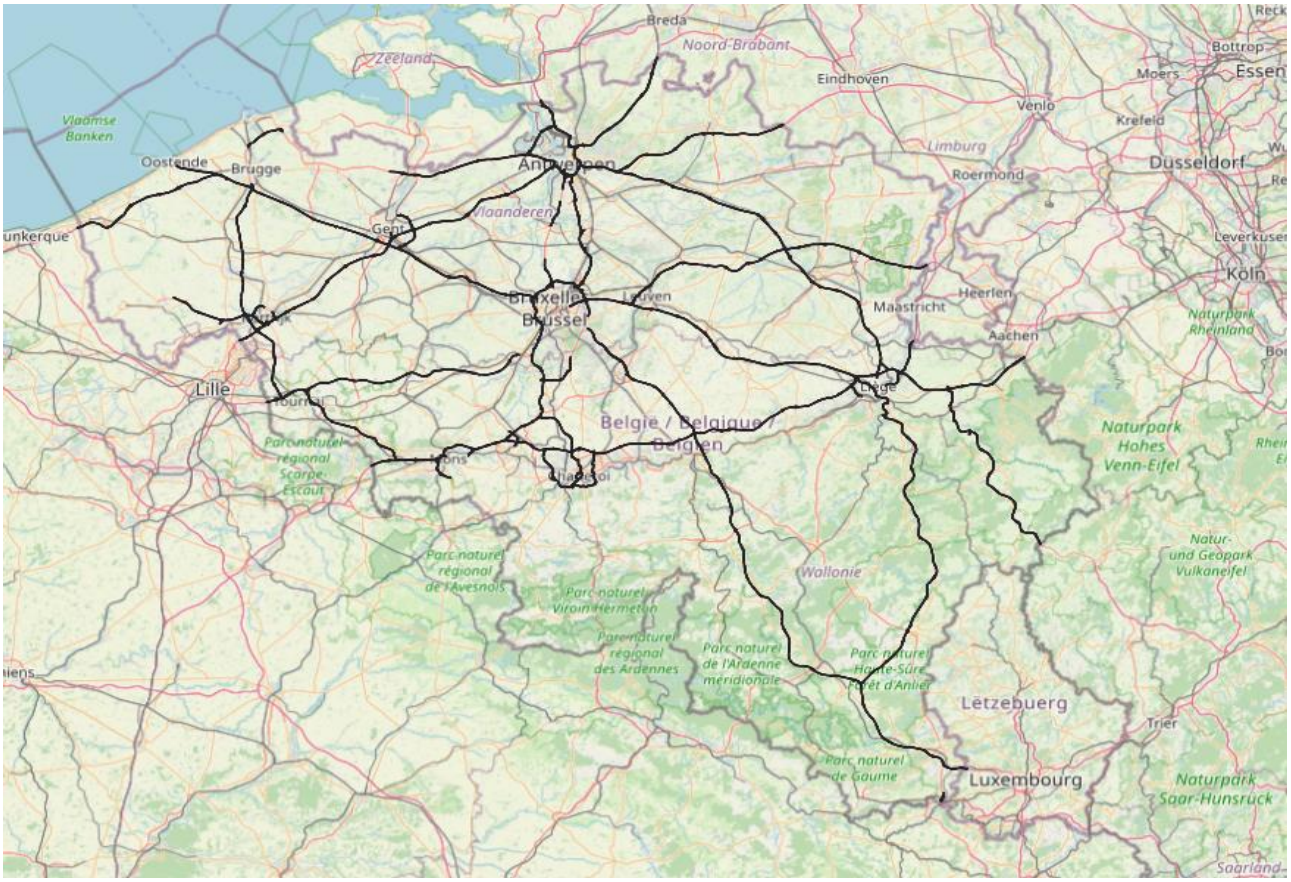

Figure A1: Motorway roads

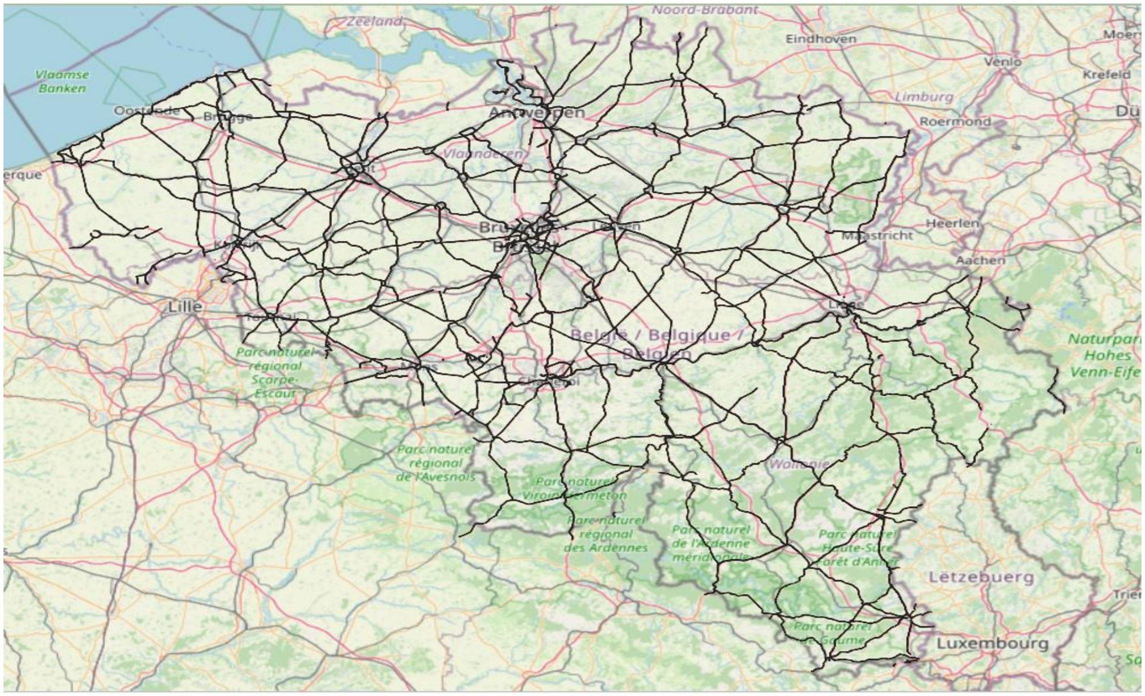

Figure A2: Primary roads

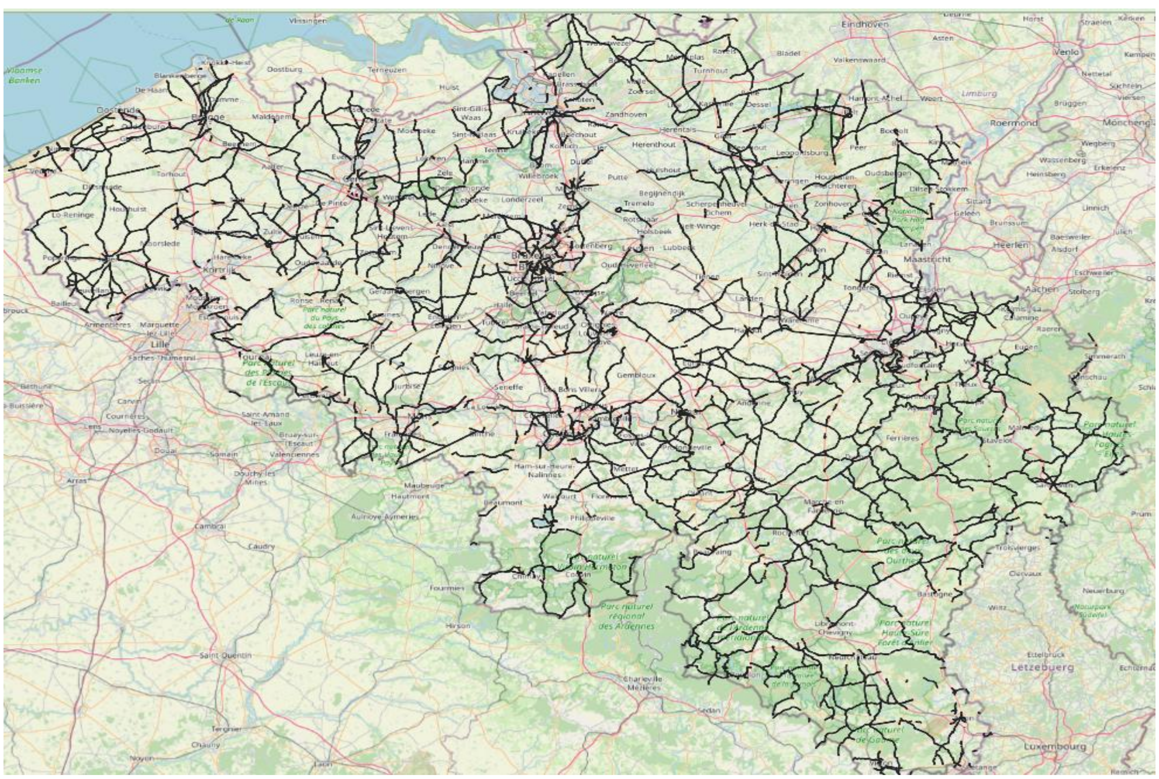

Figure A3: Secondary roads

2. Tailgating Events along Motorways, Primary and Secondary Roads

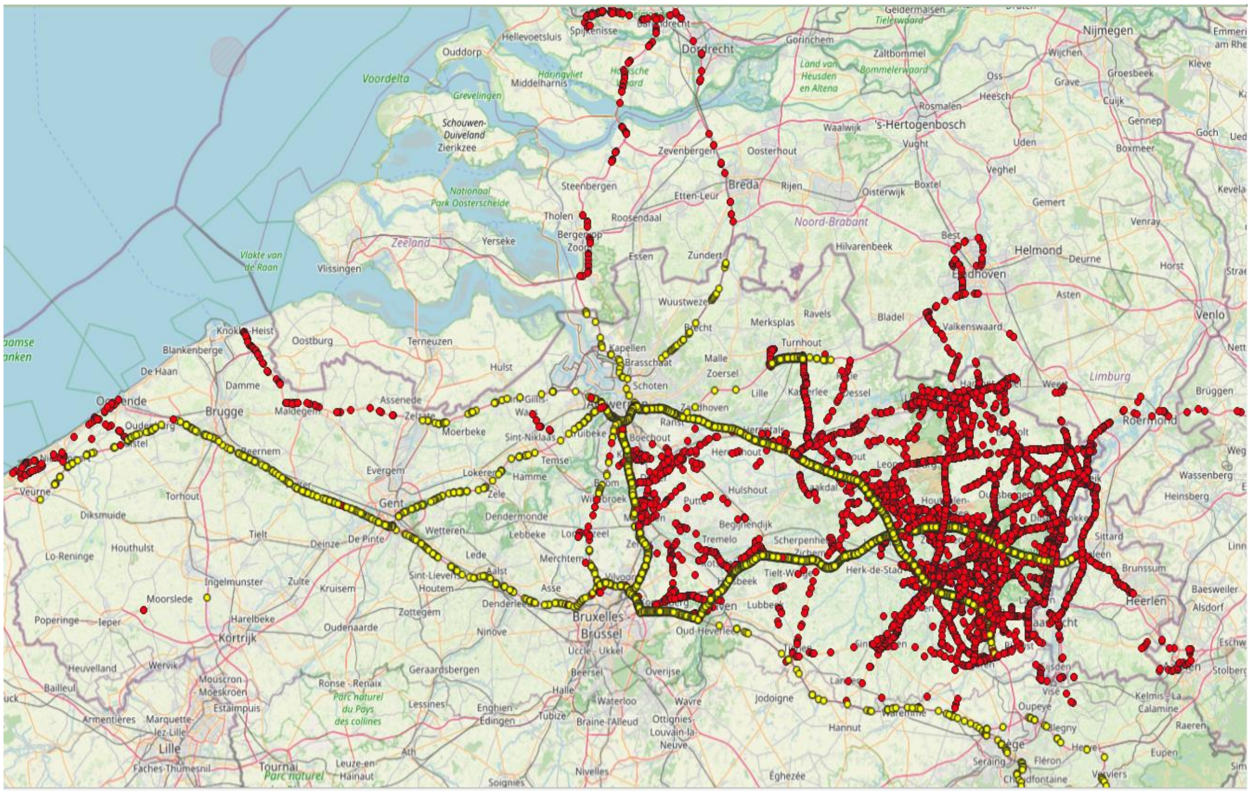

Figure A4: Tailgating events along motorways

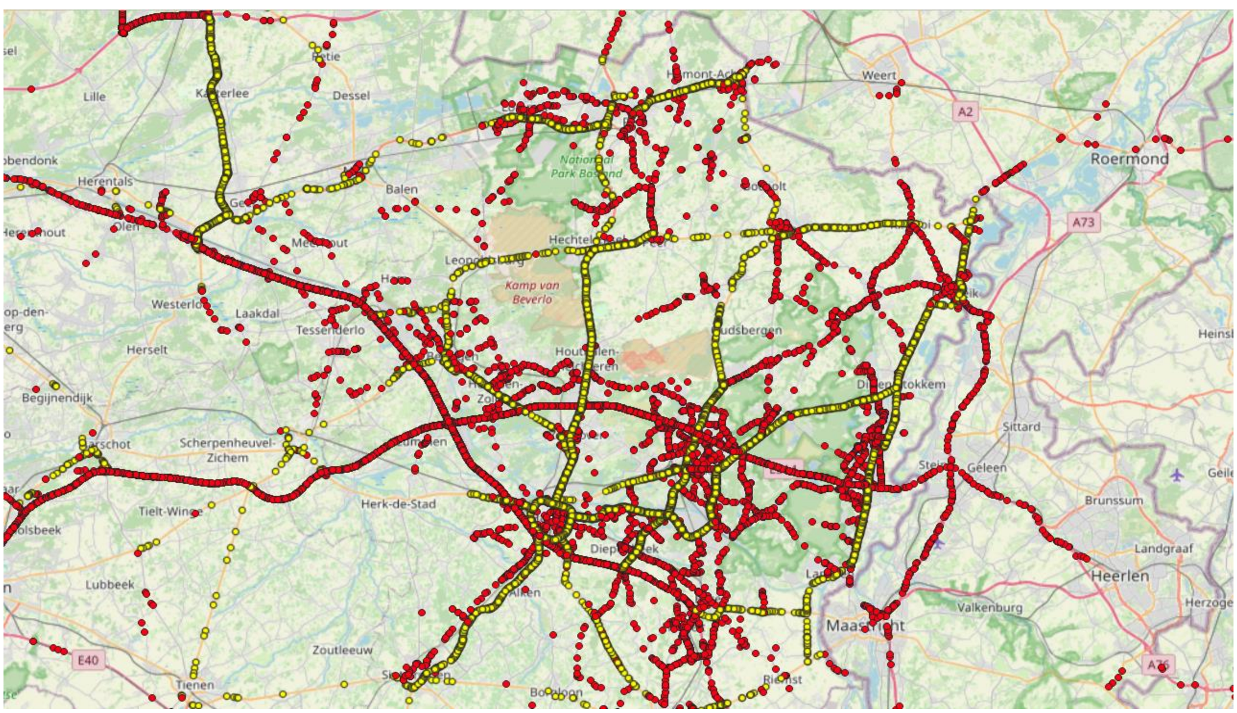

Figure A5: Tailgating events along primary roads

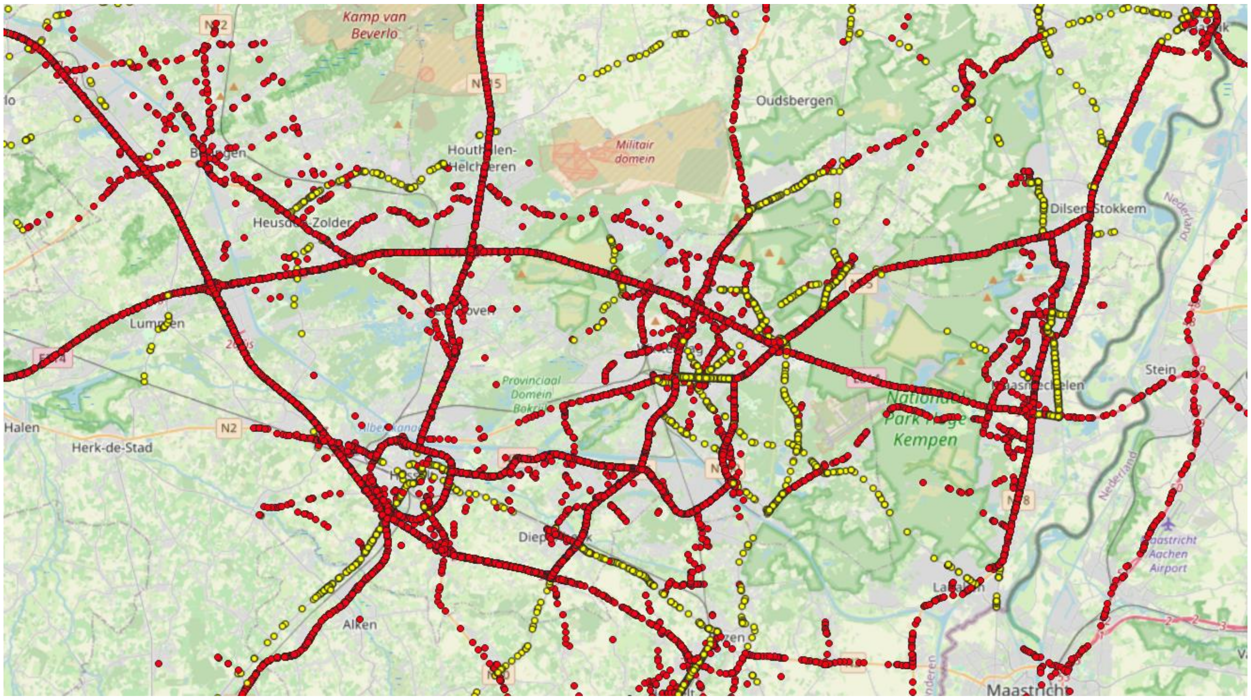

Figure A6: Tailgating events along Secondary roads

3. Speeding Events along Motorways, Primary and Secondary Roads

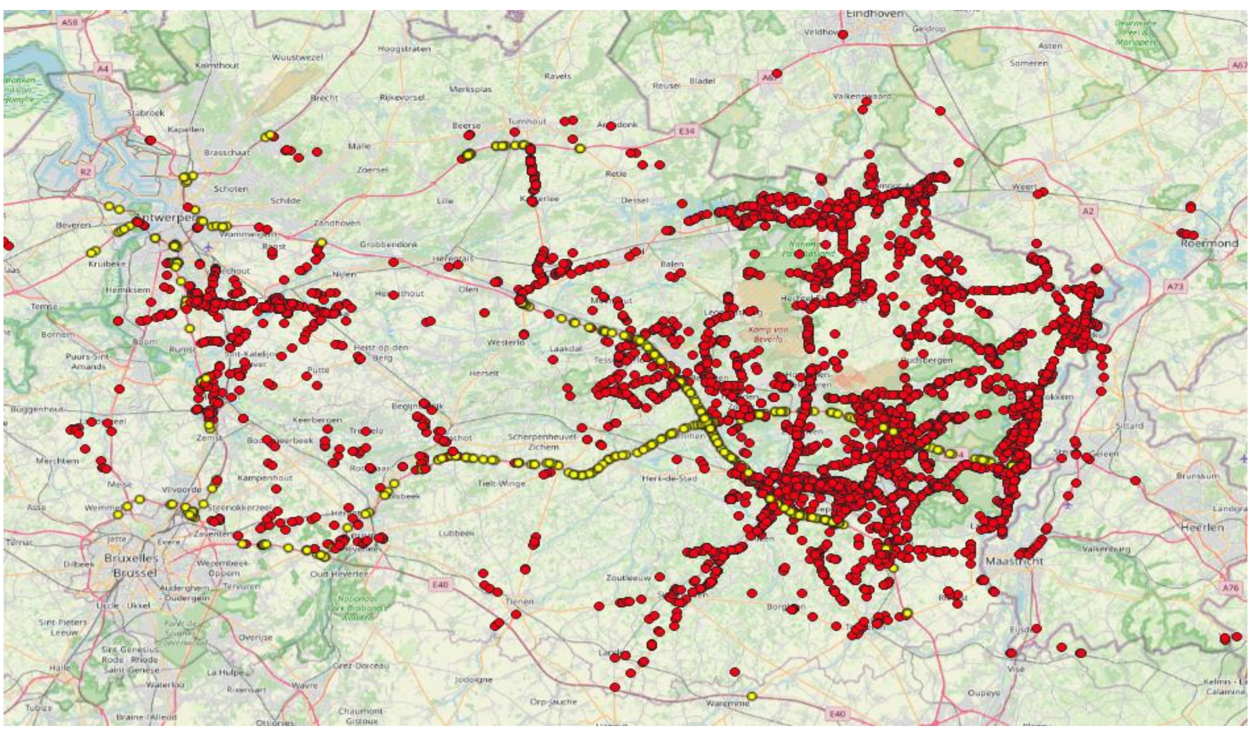

Figure A7: Speeding events along Motorways

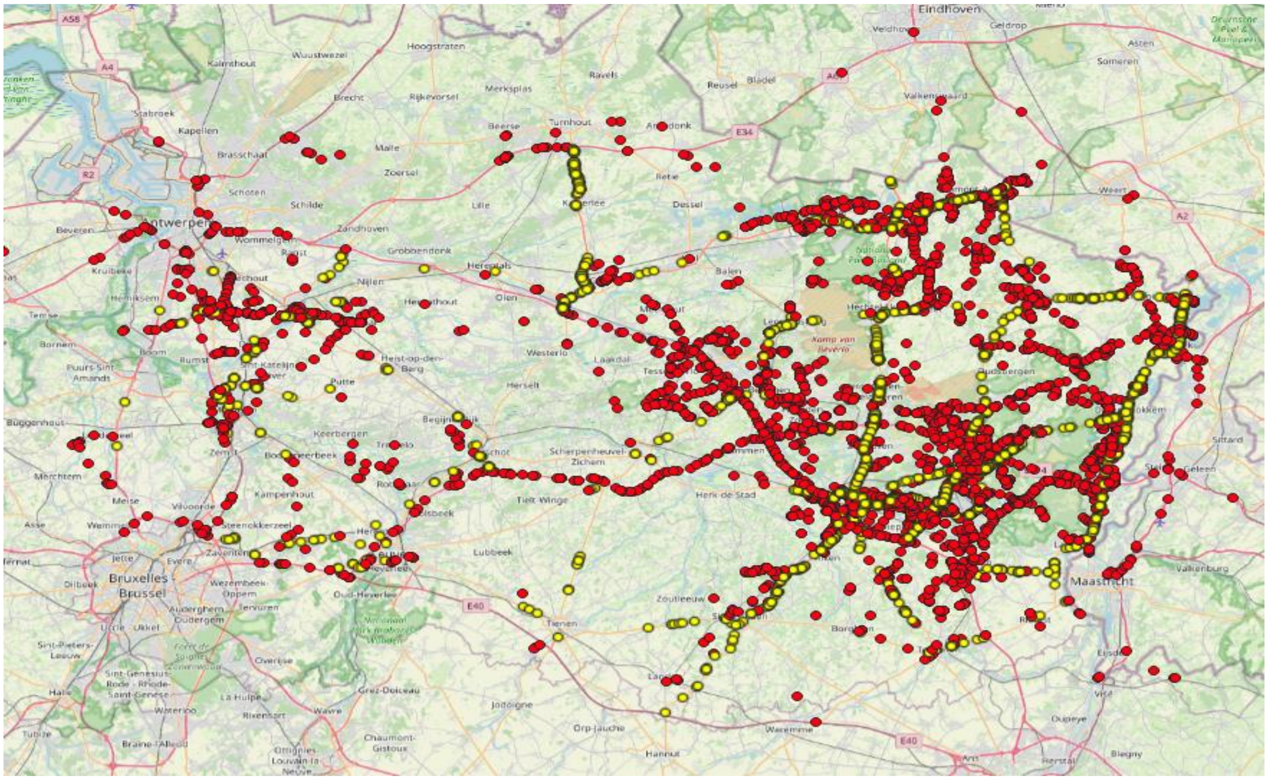

Figure A8: Speeding events along Primary roads

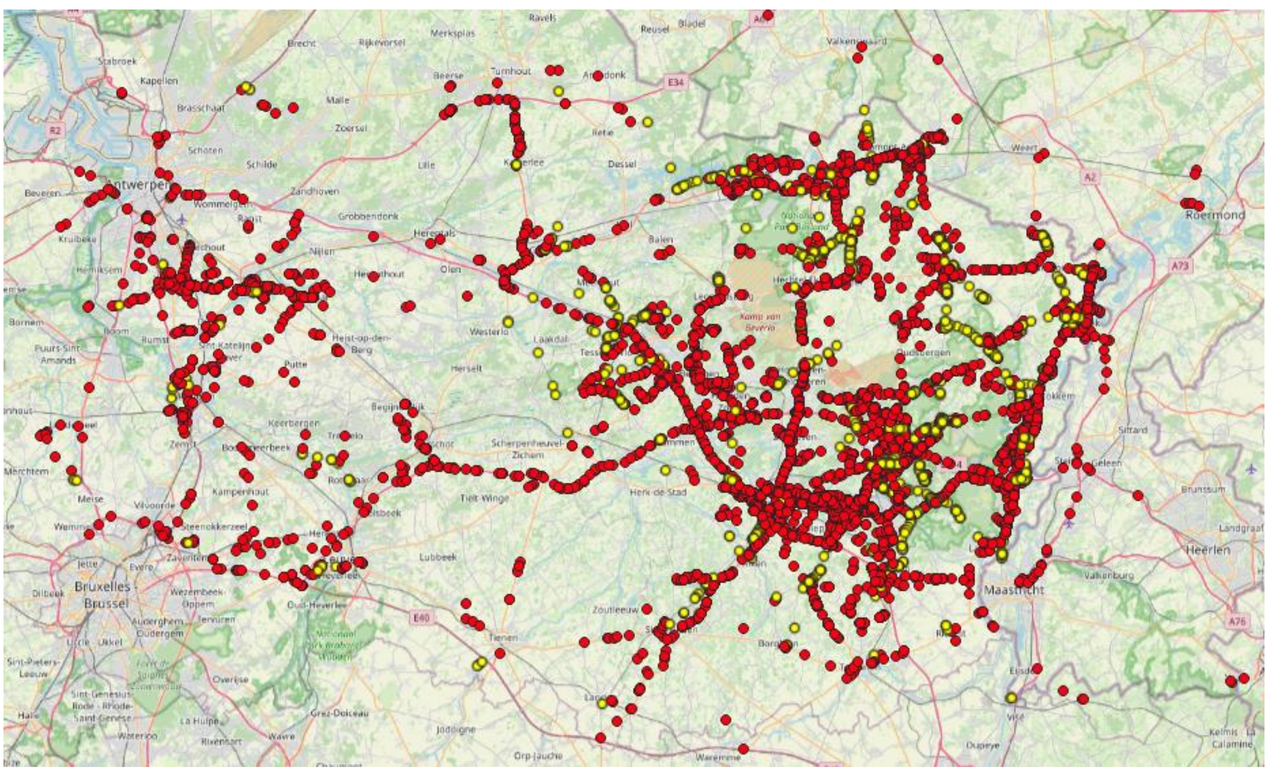

Figure A9: Speeding events along Secondary roads

APPENDIX B

SPSS Output for Speeding and Tailgating Events

B1: Checking Distance as a Covariate

1. Speeding event without distance as a covariate

Estimates

Phase	Mean	Std. Error	95% Wald Confidence Interval Lower	Upper
Phase 1	194.73	24.517	152.14	249.23
Phase 2	195.73	21.807	157.33	243.49
Phase 3	189.02	20.000	153.62	232.58
Phase 4	178.43	19.919	143.37	222.07

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	8472.500	1	<.001
Phase	.420	3	.936

Dependent Variable: Speeding_by_phases
Model: (Intercept), Phase

- Speeding event with distance as a covariate

Estimates

Phase	Mean	Std. Error	95% Wald Confidence Interval Lower	Upper
Phase 1	164.78	18.208	132.69	204.62
Phase 2	178.97	16.929	148.69	215.43
Phase 3	167.90	15.924	139.42	202.20
Phase 4	178.72	19.642	144.09	221.68

Covariates appearing in the model are fixed at the following values: Distance=929.4577

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	1599.380	1	<.001
Phase	6.711	3	.082
Phase * Distance	6.898	3	.075
Distance	40.202	1	<.001

Dependent Variable: Speeding_by_phases
Model: (Intercept), Phase, Phase * Distance, Distance

- Tailgating without distance as a covariate

Estimates

Phase	Mean	Std. Error	95% Wald Confidence Interval Lower	Upper
Phase 1	654.45	116.992	461.01	929.05
Phase 2	587.65	72.962	460.71	749.55
Phase 3	641.57	88.817	489.11	841.55
Phase 4	523.24	66.672	407.60	671.68

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	7907.516	1	<.001
Phase	1.592	3	.661

Dependent Variable: Tailgating_by_phases
Model: (Intercept), Phase

- Tailgating with distance as a covariate

Estimates

Phase	Mean	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper
Phase 1	437.75	40.226	365.60	524.14
Phase 2	482.96	34.460	419.93	555.46
Phase 3	488.24	45.180	407.26	585.34
Phase 4	471.49	55.511	374.33	593.86

Covariates appearing in the model are fixed at the following values: Distance=929.4577

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	2285.018	1	<.001
Phase	1.874	3	.599
Phase * Distance	3.038	3	.386
Distance	127.633	1	<.001

Dependent Variable: Tailgating_by_phases
Model: (Intercept), Phase, Phase * Distance, Distance

B2: Analysis by Event Severity

1. Speeding by event severity

Categorical Variable Information

			N	Percent
Factor	Group	High_Severity	204	50.0%
		Medium_Severity	204	50.0%
		Total	408	100.0%
	Phases	1	102	25.0%
		2	102	25.0%
		3	102	25.0%
		4	102	25.0%
		Total	408	100.0%

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	11428.036	1	<.001
Group	113.567	1	<.001
Phases	1.046	3	.790
Group * Phases	.177	3	.981

Dependent Variable: by_Severity_Speeding
Model: (Intercept), Group, Phases, Group * Phases

Pairwise Comparisons

(I) Phases	(J) Phases	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
1	2	.24	10.775	1	.982	-20.88	21.36
	3	3.20	10.430	1	.759	-17.24	23.65
	4	9.03	10.567	1	.393	-11.68	29.74
2	1	-.24	10.775	1	.982	-21.36	20.88
	3	2.97	9.569	1	.757	-15.79	21.72
	4	8.79	9.718	1	.366	-10.26	27.84
3	1	-3.20	10.430	1	.759	-23.65	17.24
	2	-2.97	9.569	1	.757	-21.72	15.79
	4	5.82	9.334	1	.533	-12.47	24.12
4	1	-9.03	10.567	1	.393	-29.74	11.68
	2	-8.79	9.718	1	.366	-27.84	10.26
	3	-5.82	9.334	1	.533	-24.12	12.47

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable by_Severity_Speeding

Pairwise Comparisons

(I) Group*Phases			Difference (I-J)	Std. Error	df	Sig.	Confidence Interval Lower	Confidence Interval Upper
High Severity	1	2	-3.37	24.413	1	0.890	-51.21	44.48
		3	1.67	23.355	1	0.943	-44.10	47.45
		4	15.88	23.078	1	0.491	-29.35	61.12
	2	1	3.37	24.413	1	0.890	-44.48	51.21
		3	5.04	21.831	1	0.817	-37.75	47.83
		4	19.25	21.535	1	0.380	-22.96	61.46
	3	1	-1.67	23.355	1	0.943	-47.45	44.10
		2	-5.04	21.831	1	0.817	-47.83	37.75
		4	14.21	20.327	1	0.484	-25.63	54.05
	4	1	-15.88	23.078	1	0.491	-61.12	29.35
		2	-19.25	21.535	1	0.371	-61.46	22.96
		3	-14.21	20.327	1	0.484	-54.05	25.63
Medium Severity	1	2	0.52	9.264	1	0.955	-17.64	18.68
		3	2.19	9.089	1	0.809	-15.62	20.01
		4	4.29	9.481	1	0.651	-14.29	22.87
	2	1	-0.52	9.264	1	0.955	-18.68	17.64
		3	1.67	8.099	1	0.836	-14.20	17.55
		4	3.77	8.536	1	0.659	-12.96	20.50
	3	1	-2.19	9.089	1	0.809	-20.01	15.62
		2	-1.67	8.099	1	0.836	-17.55	14.20
		4	2.10	8.346	1	0.802	-14.26	18.45
	4	1	-4.29	9.481	1	0.651	-22.87	14.29
		2	-3.77	8.536	1	0.659	-20.50	12.96
		3	-2.10	8.346	1	0.802	-18.45	14.26

Pairwise comparisons of estimated marginal means based on the original scale of dependent

a. The mean difference is significant at the .05 level.

2. Tailgating by event severity

Categorical Variable Information

			N	Percent
Factor	Group	High_Severity	192	48.5%
		Medium_Severity	204	51.5%
		Total	396	100.0%
	Phases	Phase 1	99	25.0%
		Phase 2	99	25.0%
		Phase 3	99	25.0%
		Phase 4	99	25.0%
		Total	396	100.0%

Tests of Model Effects

	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	7311.847	1	<.001
Group	141.289	1	<.001
Phases	1.804	3	.614
Group * Phases	.033	3	.998

Dependent Variable: by_severity_Tailgating
Model: (Intercept), Group, Phases, Group * Phases

Pairwise Comparisons

(I) Phases	(J) Phases	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Phase 1	Phase 2	15.76	48.935	1	.747	-80.15	111.67
	Phase 3	1.32	51.465	1	.980	-99.55	102.19
	Phase 4	43.22	46.268	1	.350	-47.47	133.90
Phase 2	Phase 1	-15.76	48.935	1	.747	-111.67	80.15
	Phase 3	-14.44	40.289	1	.720	-93.41	64.52
	Phase 4	27.46	33.396	1	.411	-38.00	92.91
Phase 3	Phase 1	-1.32	51.465	1	.980	-102.19	99.55
	Phase 2	14.44	40.289	1	.720	-64.52	93.41
	Phase 4	41.90	37.004	1	.258	-30.63	114.42
Phase 4	Phase 1	-43.22	46.268	1	.350	-133.90	47.47
	Phase 2	-27.46	33.396	1	.411	-92.91	38.00
	Phase 3	-41.90	37.004	1	.258	-114.42	30.63

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable by_severity_Tailgating

Pairwise Comparisons								
(I) Group*Phases			Difference (I-J)	Std. Error	df	Sig.	Confidence Interval Lower	Upper
High Severity	1	2	11.42	39.872	1	0.775	-66.73	89.56
		3	3.69	41.876	1	0.930	-78.39	85.76
		4	23.71	37.798	1	0.531	-50.37	97.79
	2	1	-11.42	39.872	1	0.775	-89.56	66.73
		3	-7.73	31.355	1	0.805	-69.18	53.72
		4	12.29	25.656	1	0.632	-37.99	62.58
	3	1	-3.69	41.876	1	0.930	-85.76	78.39
		2	7.73	31.355	1	0.805	-53.72	69.18
		4	20.02	28.671	1	0.485	-36.17	76.22
	4	1	-23.71	37.798	1	0.531	-97.79	50.37
		2	-12.29	25.656	1	0.632	-62.58	37.99
		3	-20.02	28.671	1	0.485	-76.22	36.17
Medium Severity	1	2	15.47	106.141	1	0.884	-192.56	223.50
		3	-10.90	112.107	1	0.923	-230.63	208.82
		4	76.73	100.293	1	0.444	-119.85	273.30
	2	1	-15.47	106.141	1	0.884	-223.50	192.56
		3	-26.37	94.932	1	0.781	-212.44	159.69
		4	61.25	80.641	1	0.447	-96.80	219.31
	3	1	10.90	112.107	1	0.923	-208.82	230.63
		2	26.37	94.932	1	0.781	-159.69	212.44
		4	87.63	88.346	1	0.321	-85.53	260.78
	4	1	-76.73	100.293	1	0.444	-273.30	119.85
		2	-61.25	80.641	1	0.447	-219.31	96.80
		3	-87.63	88.346	1	0.321	-260.78	85.53

Pairwise comparisons of estimated marginal means based on the original scale of dependent

a. The mean difference is significant at the .05 level.

B3: Analysis by Risk Zone

1. Tailgating events by risk zone

Model Information

Dependent Variable	Tailgating_by_Risk_zone
Probability Distribution	Negative binomial (1)
Link Function	Log

Case Processing Summary

	N	Percent
Included	384	100.0%
Excluded	0	0.0%
Total	384	100.0%

Categorical Variable Information

			N	Percent
Factor	Group	1	180	46.9%
		3	204	53.1%
		Total	384	100.0%
	Phase	Phase 1	96	25.0%
		Phase 2	96	25.0%
		Phase 3	96	25.0%
		Phase 4	96	25.0%
		Total	384	100.0%

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	7500.540	1	<.001
Group	14.674	1	<.001
Phase	28.062	3	<.001
Group * Phase	11.342	3	.010

Dependent Variable: Tailgating_by_Risk_zone
Model: (Intercept), Group, Phase, Group * Phase

Pairwise Comparisons

(I) Phase	(J) Phase	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
Phase 1	Phase 2	139.38[a]	39.043	1	<.001	62.86	215.90
	Phase 3	133.69[a]	39.990	1	<.001	55.31	212.07
	Phase 4	166.01[a]	37.457	1	<.001	92.59	239.42
Phase 2	Phase 1	-139.38[a]	39.043	1	<.001	-215.90	-62.86
	Phase 3	-5.68	28.160	1	.840	-60.88	49.51
	Phase 4	26.63	24.429	1	.276	-21.25	74.51
Phase 3	Phase 1	-133.69[a]	39.990	1	<.001	-212.07	-55.31
	Phase 2	5.68	28.160	1	.840	-49.51	60.88
	Phase 4	32.31	25.916	1	.212	-18.48	83.11
Phase 4	Phase 1	-166.01[a]	37.457	1	<.001	-239.42	-92.59
	Phase 2	-26.63	24.429	1	.276	-74.51	21.25
	Phase 3	-32.31	25.916	1	.212	-83.11	18.48

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Tailgating_by_Risk_zone

a. The mean difference is significant at the .05 level.

Pairwise Comparisons								
(I) Group*Phase			Difference (I-J)	Std. Error	df	Sig.	Confidence Interval Lower	Confidence Interval Upper
High Risk Zone	1	2	217.38[a]	57.434	1	0.0002	104.81	329.95
		3	220.04[a]	56.465	1	0.0001	109.37	330.71
		4	236.38[a]	55.351	1	0.0000	127.89	344.86
	2	1	-217.38[a]	57.434	1	0.0002	-329.95	-104.81
		3	2.67	32.090	1	0.9338	-60.23	65.56
		4	19.00	30.086	1	0.5277	-39.97	77.97
	3	1	-220.04[a]	56.465	1	0.0001	-330.71	-109.37
		2	-2.67	32.090	1	0.9338	-65.56	60.23
		4	16.33	28.193	1	0.5624	-38.92	71.59
	4	1	-236.38[a]	55.351	1	0.0000	-344.86	-127.89
		2	-19.00	30.086	1	0.5277	-77.97	39.97
		3	-16.33	28.193	1	0.5624	-71.59	38.92
Low Risk Zone	1	2	48.22	51.456	1	0.3487	-52.64	149.07
		3	26.63	59.925	1	0.6568	-90.82	144.08
		4	85.53	50.805	1	0.0923	-14.05	185.10
	2	1	-48.22	51.456	1	0.3487	-149.07	52.64
		3	-21.59	47.657	1	0.6506	-114.99	71.82
		4	37.31	35.514	1	0.2934	-32.29	106.92
	3	1	-26.63	59.925	1	0.6568	-144.08	90.82
		2	21.59	47.657	1	0.6506	-71.82	114.99
		4	58.90	46.952	1	0.2097	-33.12	150.93
	4	1	-85.53	50.805	1	0.0923	-185.10	14.05
		2	-37.31	35.514	1	0.2934	-106.92	32.29
		3	-58.90	46.952	1	0.2097	-150.93	33.12

Pairwise comparisons of estimated marginal means based on the original scale of dependent

a. The mean difference is significant at the .05 level.

2. Speeding events by risk zone

Model Information

Dependent Variable	Speeding_by_Risk_Zone
Probability Distribution	Negative binomial (1)
Link Function	Log

Case Processing Summary

	N	Percent
Included	372	100.0%
Excluded	0	0.0%
Total	372	100.0%

Categorical Variable Information

			N	Percent
Factor	Group	High_Risk	164	44.1%
		Low_Risk	208	55.9%
		Total	372	100.0%
	Phase	1	93	25.0%
		2	93	25.0%
		3	93	25.0%
		4	93	25.0%
		Total	372	100.0%

Estimates

				95% Wald Confidence Interval	
Group	Phase	Mean	Std. Error	Lower	Upper
High_Risk	1	33.93	6.421	23.41	49.16
	2	28.59	4.890	20.44	39.97
	3	26.27	4.932	18.18	37.95
	4	24.02	4.588	16.52	34.93
Low_Risk	1	90.08	11.139	70.69	114.78
	2	75.13	8.258	60.57	93.20
	3	74.21	7.871	60.28	91.36
	4	65.94	8.001	51.99	83.65

Tests of Model Effects

	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	4937.554	1	<.001
Group	83.776	1	<.001
Phase	4.430	3	.219
Group * Phase	.071	3	.995

Dependent Variable: Speeding_by_Risk_Zone
Model: (Intercept), Group, Phase, Group * Phase

Pairwise Comparisons

						95% Wald Confidence Interval for Difference	
(I) Phase	(J) Phase	Mean Difference (I-J)	Std. Error	df	Sig.	Lower	Upper
1	2	8.94	7.826	1	.253	-6.40	24.28
	3	11.13	7.856	1	.157	-4.27	26.53
	4	15.48[a]	7.702	1	.044	.38	30.57
2	1	-8.94	7.826	1	.253	-24.28	6.40
	3	2.19	6.698	1	.743	-10.94	15.32
	4	6.54	6.517	1	.315	-6.23	19.31
3	1	-11.13	7.856	1	.157	-26.53	4.27
	2	-2.19	6.698	1	.743	-15.32	10.94
	4	4.35	6.553	1	.507	-8.49	17.19
4	1	-15.48[a]	7.702	1	.044	-30.57	-.38
	2	-6.54	6.517	1	.315	-19.31	6.23
	3	-4.35	6.553	1	.507	-17.19	8.49

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Speeding_by_Risk_Zone

a. The mean difference is significant at the .05 level.

Pairwise Comparisons								
(I) Group*Phase			Difference (I-J)	Std. Error	df	Sig.	Confidence Interval Lower	Confidence Interval Upper
High Risk_Zone	1	2	5.34	8.071	1	0.508	-10.48	21.16
		3	7.66	8.096	1	0.344	-8.21	23.53
		4	9.90	7.891	1	0.210	-5.56	25.37
	2	1	-5.34	8.071	1	0.508	-21.16	10.48
		3	2.32	6.945	1	0.739	-11.30	15.93
		4	4.56	6.705	1	0.496	-8.58	17.70
	3	1	-7.66	8.096	1	0.344	-23.53	8.21
		2	-2.32	6.945	1	0.739	-15.93	11.30
		4	2.24	6.736	1	0.739	-10.96	15.45
	4	1	-9.90	7.891	1	0.210	-25.37	5.56
		2	-4.56	6.705	1	0.496	-17.70	8.58
		3	-2.24	6.736	1	0.739	-15.45	10.96
Low_Risk_Zone	1	2	14.94	13.866	1	0.281	-12.23	42.12
		3	15.87	13.639	1	0.245	-10.87	42.60
		4	24.13	13.714	1	0.078	-2.74	51.01
	2	1	-14.94	13.866	1	0.281	-42.12	12.23
		3	0.92	11.408	1	0.936	-21.44	23.28
		4	9.19	11.498	1	0.424	-13.34	31.73
	3	1	-15.87	13.639	1	0.245	-42.60	10.87
		2	-0.92	11.408	1	0.936	-23.28	21.44
		4	8.27	11.223	1	0.461	-13.73	30.27
	4	1	-24.13	13.714	1	0.078	-51.01	2.74
		2	-9.19	11.498	1	0.424	-31.73	13.34
		3	-8.27	11.223	1	0.461	-30.27	13.73

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable

a. The mean difference is significant at the .05 level.

B4: Analysis by Road Categories

1. **Tailgating events within the high-risk zone by road categories**

Categorical Variable Information

			N	Percent
Factor	group	MW	124	31.6%
		PR	148	37.7%
		SR	121	30.8%
		Total	393	100.0%
	Phases	Phase 1	98	24.9%
		Phase 2	111	28.2%
		Phase 3	102	26.0%
		Phase 4	82	20.9%
		Total	393	100.0%

Tests of Model Effects

Source	Type III Wald Chi-Square	df	Sig.
(Intercept)	2615.941	1	<.001
group	93.136	2	<.001
Phases	12.033	3	.007
group * Phases	6.438	6	.376

Dependent Variable: Tailgating_R.Type
Model: (Intercept), group, Phases, group * Phases

Pairwise Comparisons

(I) Phases	(J) Phases	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
Phase 1	Phase 2	23.21[a]	10.416	1	.026	2.80	43.63
	Phase 3	25.61[a]	10.128	1	.011	5.76	45.46
	Phase 4	29.60[a]	10.169	1	.004	9.66	49.53
Phase 2	Phase 1	-23.21[a]	10.416	1	.026	-43.63	-2.80
	Phase 3	2.40	6.577	1	.716	-10.49	15.29
	Phase 4	6.38	6.641	1	.337	-6.63	19.40
Phase 3	Phase 1	-25.61[a]	10.128	1	.011	-45.46	-5.76
	Phase 2	-2.40	6.577	1	.716	-15.29	10.49
	Phase 4	3.98	6.179	1	.519	-8.13	16.09
Phase 4	Phase 1	-29.60[a]	10.169	1	.004	-49.53	-9.66
	Phase 2	-6.38	6.641	1	.337	-19.40	6.63
	Phase 3	-3.98	6.179	1	.519	-16.09	8.13

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Tailgating_R.Type

a. The mean difference is significant at the .05 level.

Pairwise Comparisons

Group*Phases			Difference (I-J)	Std. Error	df	Sig.	95% Wald Lower	95% Wald Upper
MW	1	2	55.27[a]	26.530	1	0.037	3.27	107.27
		3	55.90[a]	26.245	1	0.033	4.46	107.34
		4	56.85[a]	26.629	1	0.033	4.66	109.04
	2	1	-55.27[a]	26.530	1	0.037	-107.27	-3.27
		3	0.63	8.780	1	0.943	-16.58	17.84
		4	1.58	9.868	1	0.873	-17.76	20.92
	3	1	-55.90[a]	26.245	1	0.033	-107.34	-4.46
		2	-0.63	8.780	1	0.943	-17.84	16.58
		4	0.95	9.074	1	0.917	-16.83	18.73
	4	1	-56.85[a]	26.629	1	0.033	-109.04	-4.66
		2	-1.58	9.868	1	0.873	-20.92	17.76
		3	-0.95	9.074	1	0.917	-18.73	16.83
PR	1	2	8.90	30.728	1	0.772	-51.33	69.12
		3	25.50	29.074	1	0.381	-31.49	82.48
		4	21.44	31.724	1	0.499	-40.73	83.62
	2	1	-8.90	30.728	1	0.772	-69.12	51.33
		3	16.60	24.982	1	0.506	-32.37	65.56
		4	12.55	28.022	1	0.654	-42.38	67.47
	3	1	-25.50	29.074	1	0.381	-82.48	31.49
		2	-16.60	24.982	1	0.506	-65.56	32.37
		4	-4.05	26.198	1	0.877	-55.40	47.30
	4	1	-21.44	31.724	1	0.499	-83.62	40.73
		2	-12.55	28.022	1	0.654	-67.47	42.38
		3	4.05	26.198	1	0.877	-47.30	55.40
SR	1	2	5.41	7.695	1	0.482	-9.67	20.49
		3	5.17	7.631	1	0.498	-9.79	20.13
		4	11.57	6.718	1	0.085	-1.59	24.74
	2	1	-5.41	7.695	1	0.482	-20.49	9.67
		3	-0.24	6.598	1	0.971	-13.17	12.69
		4	6.16	5.516	1	0.264	-4.65	16.97
	3	1	-5.17	7.631	1	0.498	-20.13	9.79
		2	0.24	6.598	1	0.971	-12.69	13.17
		4	6.40	5.427	1	0.238	-4.23	17.04
	4	1	-11.57	6.718	1	0.085	-24.74	1.59
		2	-6.16	5.516	1	0.264	-16.97	4.65
		3	-6.40	5.427	1	0.238	-17.04	4.23

Pairwise comparisons of estimated marginal means based on the original scale of

a. The mean difference is significant at the .05 level.

2. **Tailgating events within the high-risk zone by road categories**

Categorical Variable Information

			N	Percent
Factor	group	MW	137	26.6%
		PR	189	36.7%
		SR	189	36.7%
		Total	515	100.0%
	phase	Phase 1	139	27.0%
		Phase 2	137	26.6%
		Phase 3	130	25.2%
		Phase 4	109	21.2%
		Total	515	100.0%

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	4414.087	1	<.001
group	38.550	2	<.001
phase	3.637	3	.303
group * phase	1.636	6	.950

Dependent Variable: Tailgating_R.Type
Model: (Intercept), group, phase, group * phase

Pairwise Comparisons

(I) phase	(J) phase	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Phase 1	Phase 2	5.15	10.129	1	.611	-14.70	25.01
	Phase 3	1.63	12.392	1	.895	-22.66	25.92
	Phase 4	16.64	10.019	1	.097	-3.00	36.27
Phase 2	Phase 1	-5.15	10.129	1	.611	-25.01	14.70
	Phase 3	-3.52	10.988	1	.749	-25.06	18.01
	Phase 4	11.48	8.219	1	.162	-4.63	27.59
Phase 3	Phase 1	-1.63	12.392	1	.895	-25.92	22.66
	Phase 2	3.52	10.988	1	.749	-18.01	25.06
	Phase 4	15.01	10.887	1	.168	-6.33	36.34
Phase 4	Phase 1	-16.64	10.019	1	.097	-36.27	3.00
	Phase 2	-11.48	8.219	1	.162	-27.59	4.63
	Phase 3	-15.01	10.887	1	.168	-36.34	6.33

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Tailgating_R.Type

Pairwise Comparisons

Group*phase			Difference (I-J)	Std. Error	df	Sig.	95% Wald Lower	95% Wald Upper
MW	1	2	4.08	28.989	1	0.888	-52.74	60.89
		3	-34.42	51.371	1	0.503	-135.11	66.26
		4	12.60	31.119	1	0.686	-48.39	73.59
	2	1	-4.08	28.989	1	0.888	-60.89	52.74
		3	-38.50	47.381	1	0.417	-131.36	54.37
		4	8.52	23.966	1	0.722	-38.45	55.50
	3	1	34.42	51.371	1	0.503	-66.26	135.11
		2	38.50	47.381	1	0.417	-54.37	131.36
		4	47.02	48.714	1	0.334	-48.46	142.50
	4	1	-12.60	31.119	1	0.686	-73.59	48.39
		2	-8.52	23.966	1	0.722	-55.50	38.45
		3	-47.02	48.714	1	0.334	-142.50	48.46
PR	1	2	9.37	18.608	1	0.615	-27.10	45.84
		3	16.26	17.779	1	0.360	-18.58	51.11
		4	21.95	17.798	1	0.217	-12.93	56.83
	2	1	-9.37	18.608	1	0.615	-45.84	27.10
		3	6.90	15.246	1	0.651	-22.99	36.78
		4	12.58	15.268	1	0.410	-17.34	42.51
	3	1	-16.26	17.779	1	0.360	-51.11	18.58
		2	-6.90	15.246	1	0.651	-36.78	22.99
		4	5.69	14.247	1	0.690	-22.24	33.61
	4	1	-21.95	17.798	1	0.217	-56.83	12.93
		2	-12.58	15.268	1	0.410	-42.51	17.34
		3	-5.69	14.247	1	0.690	-33.61	22.24
SR	1	2	2.71	8.727	1	0.756	-14.39	19.82
		3	6.78	8.154	1	0.406	-9.21	22.76
		4	13.62	7.621	1	0.074	-1.32	28.55
	2	1	-2.71	8.727	1	0.756	-19.82	14.39
		3	4.06	8.099	1	0.616	-11.81	19.94
		4	10.90	7.562	1	0.149	-3.92	25.72
	3	1	-6.78	8.154	1	0.406	-22.76	9.21
		2	-4.06	8.099	1	0.616	-19.94	11.81
		4	6.84	6.893	1	0.321	-6.67	20.35
	4	1	-13.62	7.621	1	0.074	-28.55	1.32
		2	-10.90	7.562	1	0.149	-25.72	3.92
		3	-6.84	6.893	1	0.321	-20.35	6.67

Pairwise comparisons of estimated marginal means based on the original scale of

a. The mean difference is significant at the .05 level.

3. Speeding events by road categories

Categorical Variable Information

			N	Percent
Factor	Group	Motorway	156	28.3%
		Primary roads	200	36.2%
		Secondary roads	196	35.5%
		Total	552	100.0%
	Phases	1	138	25.0%
		2	138	25.0%
		3	138	25.0%
		4	138	25.0%
		Total	552	100.0%

Tests of Model Effects

	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	4072.129	1	<.001
Group	80.076	2	<.001
Phases	1.752	3	.625
Group * Phases	.735	6	.994

Dependent Variable: Speeding_by_road_types
Model: (Intercept), Group, Phases, Group * Phases

Pairwise Comparisons

(I) Phases	(J) Phases	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
1	2	2.09	4.997	1	.676	-7.71	11.88
	3	.53	5.098	1	.917	-9.46	10.53
	4	5.48	5.050	1	.278	-4.42	15.38
2	1	-2.09	4.997	1	.676	-11.88	7.71
	3	-1.55	4.123	1	.706	-9.64	6.53
	4	3.39	4.064	1	.404	-4.57	11.35
3	1	-.53	5.098	1	.917	-10.53	9.46
	2	1.55	4.123	1	.706	-6.53	9.64
	4	4.94	4.187	1	.238	-3.26	13.15
4	1	-5.48	5.050	1	.278	-15.38	4.42
	2	-3.39	4.064	1	.404	-11.35	4.57
	3	-4.94	4.187	1	.238	-13.15	3.26

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Speeding_by_road_types

Pairwise Comparisons								
			Difference (I-J)	Std. Error	df	Sig.	Confidence Interval	
Group*Phases							Lower	Upper
Motorways	1	2	2.59	5.302	1	0.625	-7.80	12.98
		3	0.54	5.679	1	0.924	-10.59	11.67
		4	2.90	5.929	1	0.625	-8.72	14.52
	2	1	-2.59	5.302	1	0.625	-12.98	7.80
		3	-2.05	3.808	1	0.590	-9.51	5.41
		4	0.31	4.172	1	0.941	-7.87	8.48
	3	1	-0.54	5.679	1	0.924	-11.67	10.59
		2	2.05	3.808	1	0.590	-5.41	9.51
		4	2.36	4.641	1	0.611	-6.74	11.46
	4	1	-2.90	5.929	1	0.625	-14.52	8.72
		2	-0.31	4.172	1	0.941	-8.48	7.87
		3	-2.36	4.641	1	0.611	-11.46	6.74
Primary roads	1	2	4.06	15.221	1	0.790	-25.77	33.89
		3	4.68	14.171	1	0.741	-23.10	32.46
		4	9.82	14.010	1	0.483	-17.64	37.28
	2	1	-4.06	15.221	1	0.790	-33.89	25.77
		3	0.62	12.825	1	0.961	-24.52	25.76
		4	5.76	12.647	1	0.649	-19.03	30.55
	3	1	-4.68	14.171	1	0.741	-32.46	23.10
		2	-0.62	12.825	1	0.961	-25.76	24.52
		4	5.14	11.362	1	0.651	-17.13	27.41
	4	1	-9.82	14.010	1	0.483	-37.28	17.64
		2	-5.76	12.647	1	0.649	-30.55	19.03
		3	-5.14	11.362	1	0.651	-27.41	17.13
Secondary roads	1	2	-1.37	7.483	1	0.855	-16.03	13.30
		3	-2.31	7.711	1	0.765	-17.42	12.81
		4	5.78	6.601	1	0.382	-7.16	18.71
	2	1	1.37	7.483	1	0.855	-13.30	16.03
		3	-0.94	7.433	1	0.899	-15.51	13.63
		4	7.14	6.273	1	0.255	-5.15	19.44
	3	1	2.31	7.711	1	0.765	-12.81	17.42
		2	0.94	7.433	1	0.899	-13.63	15.51
		4	8.08	6.544	1	0.217	-4.74	20.91
	4	1	-5.78	6.601	1	0.382	-18.71	7.16
		2	-7.14	6.273	1	0.255	-19.44	5.15
		3	-8.08	6.544	1	0.217	-20.91	4.74

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable

a. The mean difference is significant at the .05 level.

Declaration of AI use

Generative AI was used to browse the subject, come up with new ideas, and improve self-written texts.