informatics

MDPI

Article

Risk Factors Influencing Fatal Powered Two-Wheeler At-Fault and Not-at-Fault Crashes: An Application of Spatio-Temporal Hotspot and Association Rule Mining Techniques

Reuben Tamakloe

Department of Transportation Engineering, University of Seoul, 163 Seoulsiripdae-ro Dongdaemun-gu, Seoul 02504, Republic of Korea; drtamakloe@uos.ac.kr

Abstract: Studies have explored the factors influencing the safety of PTWs; however, very little has been carried out to comprehensively investigate the factors influencing fatal PTW crashes while considering the fault status of the rider in crash hotspot areas. This study employs spatio-temporal hotspot analysis and association rule mining techniques to discover hidden associations between crash risk factors that lead to fatal PTW crashes considering the fault status of the rider at statistically significant PTW crash hotspots in South Korea from 2012 to 2017. The results indicate the presence of consecutively fatal PTW crash hotspots concentrated within Korea's densely populated capital, Seoul, and new hotspots near its periphery. According to the results, violations such as over-speeding and red-light running were critical contributory factors influencing PTW crashes at hotspots during summer and at intersections. Interestingly, while reckless riding was the main traffic violation leading to PTW rider at-fault crashes at hotspots, violations such as improper safety distance and red-light running were strongly associated with PTW rider not-at-fault crashes at hotspots. In addition, while PTW rider at-fault crashes are likely to occur during summer, PTW rider not-at-fault crashes mostly occur during spring. The findings could be used for developing targeted policies for improving PTW safety at hotspots.

Keywords: spatio-temporal hotspot analysis; powered two-wheeler; association rule mining; fatal crash; crash risk factor

check for updates

Citation: Tamakloe, R. Risk Factors Influencing Fatal Powered Two-Wheeler At-Fault and Not-at-Fault Crashes: An Application of Spatio-Temporal Hotspot and Association Rule Mining Techniques. *Informatics* **2023**, *10*, 43. https://doi.org/10.3390/informatics10020043

Academic Editor: Weitian Tong

Received: 27 February 2023
Revised: 23 April 2023
Accepted: 9 May 2023
Published: 12 May 2023

1. Introduction

1.1. Background

Road traffic injuries claim the lives of numerous young people annually, with over 1.35 million fatalities reported each year, causing injuries to nearly 50 million people [1,2]. Regrettably, vulnerable road users, comprising pedestrians, cyclists, and motorcyclists, constitute more than half of all road traffic deaths. An alarming 30% of the crash fatalities reported in 2016 involved powered two and three wheelers (PTWs), including motorcycles, e-bikes, and scooters, most of which could have been prevented [3]. The Global Status Report on Road Safety 2018 documents that motorcyclists/PTWs alone accounted for 378,000 deaths, which is more than a quarter of all road traffic fatalities [4]. PTWs have a higher propensity for fatal and severe injuries than other vehicle types [5]. For example, in Australia, PTW crash death rates are over thirty times that of car users per distance traveled. In addition, PTW users account for over 43% of all deaths in the South East Asia region [4,6]. PTWs offer immense advantages such as low parking space requirements, high maneuverability, and better fuel efficiency to a diverse group of users, and more than 75% of the 300 million PTWs worldwide are found in Asia [7,8]. However, the hazardous riding behavior associated with PTWs is responsible for the majority of accidents. It remains unclear from the literature what risk factors contribute to fatal PTW crashes, both for at-fault and not-at-fault accidents, at key crash locations.

While the majority of the existing studies have primarily focused on identifying and quantifying the effect of risk factors on PTW crash severity [9,10], others have identified contributory factors leading to PTW crashes [11] and those influencing crash-type propensity [12]. Regarding the factors associated with PTW crashes, de Rome and Senserrick (2011) identified influential critical factors using a descriptive analysis procedure [13]. Other studies applied parametric and non-parametric methods to identify critical PTW contributory factors [14,15].

Notably, none of the studies in the literature comprehensively explored the chains of risk factors influencing fatal PTW crashes that occurred at critical hotspot locations identified using spatio-temporal pattern mining procedure analysis. Regardless of the high interest in PTW safety, many fundamental questions concerning the factors leading to fatal PTW crashes at critical crash locations based on the riders' fault status remain unanswered. Although road crashes can happen anywhere at any given time, research shows that spatial and temporal dependencies among crashes exist [16], and may cause contributory factors influencing a crash to differ. In addition, risk factors leading to a collision could differ based on the fault status of the driver/rider, as highlighted in the literature [17].

To establish targeted countermeasures that can help in increasing PTW safety at critical hotspot locations, it is crucial to examine the chains of factors and patterns leading to fatal PTW crashes using a more disaggregate approach considering spatio-temporal information and the riders' fault status. Understanding the contributory factors influencing fatal PTW crashes at crash hotspots identified over time and differentiated based on the riders' fault status could provide valuable information for use by traffic safety analysts worldwide.

1.2. Summary and Critique of the Literature

The causes and severity of crashes can be described under three components: the driver, vehicle, and environment. In PTW safety, research has identified driver/human factors such as age, gender, and errors as common risk factors. Young riders are particularly at risk of PTW crashes, and male riders are highly at risk of crashes on light PTWs relative to their female counterparts [18,19]. Several reasons have been attributed to this finding, ranging from lack of experience and poor hazard perception on the part of young riders, and overconfidence on the part of men [20,21]. Driver errors such as speeding, risky maneuvers, and late observations have also been identified as critical crash risk factors in recent studies [11,22,23]. Regarding environmental factors (temporal, weather, and roadway-related factors), some studies identified that crash risk increases during the early morning hours, during the weekends, spring and summer, wider roads, and poor roadway conditions [19,23], broader shoulders and more lanes resulted in a reduced crash risk [24]. While heavy PTWs significantly decrease crash risk, sports bikes tend to double the crash risk compared to other PTW types [18].

A recent study [25] identified GNI, motorized two and three wheelers per person ratio, percentage of helmet-wearing rate, and the interaction between vehicle/person ratio and motorized two and three wheelers/person ratio as significant factors affecting PTW mortality rates. In another study [26], the impact of variables on injury severity differed significantly for helmet-wearing and non-helmet-wearing motorcyclists. Factors such as age, gender, collisions, and distracted driving contributed to injury severity. A Pakistani study [27] showed that crashes involving riders over 50 years old, collisions with heavy vehicles, and speeding increased the risk of severe or fatal injuries. Similarly, a Portuguese study [28] found that several factors, such as the PTW category, male rider, and no helmet use, increased the likelihood of severe injuries in motorcycle accidents. Finally, another study in Pakistan [29] showed that young drivers, high-speed limits, poor lighting, and shiny weather conditions worsened injury severity in three-wheeled motorized rickshaw crashes.

Research shows that the factors impacting crashes and the severity of injuries sustained are likely to differ based on fault status [30,31]. PTW riders were identified to be likely at fault in rear-end crashes when there is no passenger when the passenger is male or is of the same age as the rider. In addition, head-on crashes occurring during the day are likely the fault of PTW riders. In contrast, PTW riders are likely not at fault for sideswipes during daylight conditions [32]. Regarding injury severity, a recent study separately modeled at-fault and not-at-fault crash injury severities of all vehicle crashes, including motorcycles in North Carolina, and identified that, whereas road geometry and weather features had a similar effect on the severity of both drivers, the driver's age, gender, number and types of violations, and vehicle type played critical roles in the severity of injuries sustained by not-at-fault drivers [33]. The authors acknowledged that motorcycle riders are the most vulnerable road users in both at- and not-at-fault crashes. Rezapour et al. recently analyzed PTW at-fault crash data and demonstrated that while age (>35), alcohol involvement, rural area, and speed increased the severity of crashes, wet road surfaces were associated with a reduced injury severity propensity for at-fault riders [34]. However, these studies did not provide detailed insights into how variables affected PTW crashes differently regarding the rider's fault status.

Recent studies have shown that crash hotspots will likely change in time and space. For example, a study exploring crash severity during the COVID-19 pandemic revealed that hotspots would likely change from higher-income to lower-income areas [35]. Further, as identified from recent studies that the dynamics in crash clusters or hotspots vary over time and space [16,36,37], and there could be some disparities among the factors influencing fatal PTW crashes based on when and where the crash occurred. In the area of motorcycle safety, Jiang et al. are the first to use spatial analysis techniques to identify and visualize motorcycle crash hotspots in Victoria, Australia [38]. The authors identified that motorcycle crashes are concentrated around metropolitan areas. It is noteworthy that the researchers only considered the spatial context in identifying crash hotspots instead of combining it with the temporal context of the crashes, whose results could be informative in the formulation of policies. In addition, the authors did not study and identify the critical risk factors influencing fatal crashes in hotspot areas. As crash hotspots are known to be associated with higher crash frequency and fatality, which changes from time to time, identifying crash hotspots over space and time and exploring the associations among fatal PTW crash risk factors could provide valuable insights that could be used for decision-making. The literature reviewed in this section is summarized in Table 1 below.

Table 1. Summary of the literature reviewed in this study.

Study Objective	Location	Temporal Scope	Methods Adopted	Key Findings
Examining fatal and severe life-threatening accidents at intersections involving a powered two wheeler and another vehicle [11].	Six countries (UK, The Netherlands, France, Poland, Italy, and Greece)	2015 to 2016 (n = 92)	Causation analysis	Most vehicles do not yield to the PTW, and causal chains indicate that "looked but failed to see" is still a problem in this type of collision.
Recognizing segments of motorcycle riders at a notably elevated risk of accidents and pinpointing the contributing risk factors involved [18].	Norway	2005 to 2008 (n = 3356)	Survey-based research	The majority of deadly incidents involving sports motorcycles are due to speeding. In Norway, young age, limited experience, hazardous actions, and an "unsafe" mentality appear to be particularly powerful risk elements for motorcyclists.
Analyzing the factors contributing to motorcycle crashes using a safe systems approach by utilizing the case-series data collected from a recent study that employed a case–control methodology [22].	Victoria, Australia	January 2012 and August 2014 (n = 235)	Survey-based research	Although PTW crashes, resulting in injury, involve a complex array of factors, several noteworthy connections exist between the primary contributing factor (rider or other road users) and secondary factors such as rider age, traffic density, speed, and road design issues.
To gain a more profound comprehension of how these incidents happen [23].	Bogota, Colombia	2009 (n = 400)	Case study analysis	Various factors, including the absence of clear road markings, a complex intersection, a wide road, and an inexperienced motorcyclist, combine to contribute to the incidence of accidents of this nature.
Validating a previous model to consider the geometric features of intersections when predicting the safety of motorcyclists [24].	Italy	2001 to 2006	Experimental investigation	The speed at which vehicles approach an intersection and how the intersection is designed are important factors that can help explain why motorcycle accidents occur at junctions.
Identifying the main risk factors linked with the severity of injuries sustained by motorcyclists in Rawalpindi, Pakistan [27].	Rawalpindi, Pakistan	2017 to 2019	Random parameters logit model with heterogeneity in means and variances	The likelihood of severe and deadly injuries is higher for crashes that happen on weekdays, involve riders over the age of 50, involve a collision between a motorcycle and a passenger car or a heavy vehicle, have a female passenger on the back of the motorcycle, and result from exceeding the speed limit.
Identifying the factors that increase the severity of injuries sustained by powered two-wheeler (PTW) riders in road accidents in Portugal [28].	Portugal	2010 to 2015 (n = 37,769)	Ordered logistic regression	Several factors can lead to more severe injuries, including riding a motorcycle in the PTW category, having rest days, driving on clean and dry roads between 20 h and 5 h and 59 min, traveling in rural areas with bent roads and national roads, being a male rider without a helmet, having a blood alcohol content between 0.5 g/L and 0.8 g/L, and being involved in an accident with a truck or other vehicles where the driver is injured.

Table 1. *Cont.*

Study Objective	Location	Temporal Scope	Methods Adopted	Key Findings
To determine the primary elements responsible for the error of bikers engaged in accidents [32].	Iran	2009 to 2012 (n = 90,418)	Classification and regression tree algorithm (CART)	The type of collision is the primary factor determining the probability of motorcyclists being at fault. Based on this information, the chances of a rear-end collision are the highest, while the chances of a side collision are the lowest.
To analyze and contrast how certain factors impact the degree of harm suffered by drivers involved in accidents, regardless of whether they were responsible for the incident or not [33].	North Carolina	2009 to 2013 (n = 349,454)	Proportional odds model	The impact of road characteristics, weather conditions, and geometric characteristics on crash injury severity was similar for both at-fault and not-at-fault drivers. However, the age of the driver, physical condition, gender, vehicle type, and the number and type of traffic rule violations were identified as significant factors that affect the injury severity of not-at-fault drivers compared to at-fault drivers involved in the crash.
Forecasting the extent of injury in a motorcycle accident caused by the rider at fault [34].	Wyoming	2007 to 2016 (n = 1210)	Binary logistic regression and classification trees (CT)	A number of factors that were recognized as similar by both approaches include speed limit displayed on signs, age, functional class of the highway, and adherence to speed regulations.
To examine how factors contributing to the severity of injuries sustained in alcohol/drug-impaired car crashes vary throughout the day and over time, during three specific phases of crash cycles that occurred after the Great Recession [37].	North Carolina	2008 to 2017	Random parameters logit models with heterogeneity in the means and variances	Significant temporal instability of risk factor impacts was identified.
To analyze how the COVID-19 pandemic and resulting mobility alterations have affected road traffic safety [35].	Los Angeles and New York	March, 2020	Change-point detection and difference-in-differences analysis	The areas where accidents occur have frequently changed in terms of both location and time.
To determine the crucial elements associated with the severity of motorcycle injuries [38].	Victoria, Australia	2006 to 2017 (n = 24,680)	Association rule mining and hotspot analysis	Collisions with trucks, overtaking, overspeeding, collisions late at night/early morning, and collisions with fixed objects are critical factors influencing motorcycle safety in hotspots.

1.3. Study Objective and Contribution

As discussed, studies have explored the contributory factors of PTW crashes and the factors impacting their crash risk and severity. However, to the best of our knowledge, safety researchers have not considered the rider's fault status, together with both the spatial and temporal dimensions of PTW crashes, in identifying critical risk factors influencing PTW safety. In other words, it is not clear how the patterns of critical risk factor associations influencing fatal PTW crashes differ at different spatio-temporal hotspot segments considering the fault status of the rider. Identifying risk factors influencing PTW crashes (at-fault and not-at-fault) in spatio-temporal hotspot locations is essential as it can help discover the critical issues affecting PTW crashes at critical locations to develop targeted interventions and countermeasures. In addition, identifying these factors can help allocate resources and funding for road safety initiatives and improve PTW safety education and training programs targeting the locations that really need policy interventions. This study aims to deepen the knowledge of PTW safety by employing spatio-temporal hotspot analysis and data mining tools to identify critical PTW crash locations and explore chains of risk factors influencing them based on the fault status of the rider. This study seeks to answer the following questions:

- Where are the spatio-temporal hotspot locations of fatal PTW crashes in South Korea concentrated, and what are the trends associated with them?
- What are the chains of factors associated with PTW crashes at critical crash hotspots?
- Do the elements of the chains of risk factors influencing fatal PTW crashes at critical hotspots differ when the fault status of the PTW rider is considered?

This research is expected to bridge the gap in the literature by expanding the understanding of the factor chains or groups of risk factors simultaneously contributing to fatal PTW crashes considering the riders' fault status and the space-time hotspots—providing policymakers with information on how to address these crashes at critical crash locations.

The remainder of the paper is organized as follows. Section 2 presents a study framework that describes how this study was conducted. It also describes the case study area, the data used for this study, the spatio-temporal pattern mining, and the association rules mining technique applied in this study. Section 3 provides estimation results and a detailed discussion of the findings. Finally, the key findings, recommendations, limitations, and future research directions are provided in Section 4.

2. Materials and Methods

2.1. Study Framework

As illustrated in Figure 1, this study's framework comprises three main steps. First, the data for this study are obtained, and the x-y coordinates are imported into the ArcGIS Pro software after cleaning the data and removing those with missing coordinate information. Second, all the crashes are visualized in ArcGIS. Using the base map of South Korea as the boundaries, the ArcGIS Pro software was then used to develop a spatio-temporal hotspot map based on the frequency of fatal PTW crashes. Crashes at hotspot locations are then extracted for analysis. Further, the spatio-temporal trends in the PTW crashes are examined.

A data mining technique is employed to explore the chains of factors influencing fatal PTW crashes for all PTW crashes at critical PTW hotspot locations. The data mining tool is again used to investigate the contributory factors leading to fatal PTW crashes based on the fault status of the rider. Finally, the study results are thoroughly examined, and relevant policies and countermeasures are proposed for use by policymakers both locally and internationally.

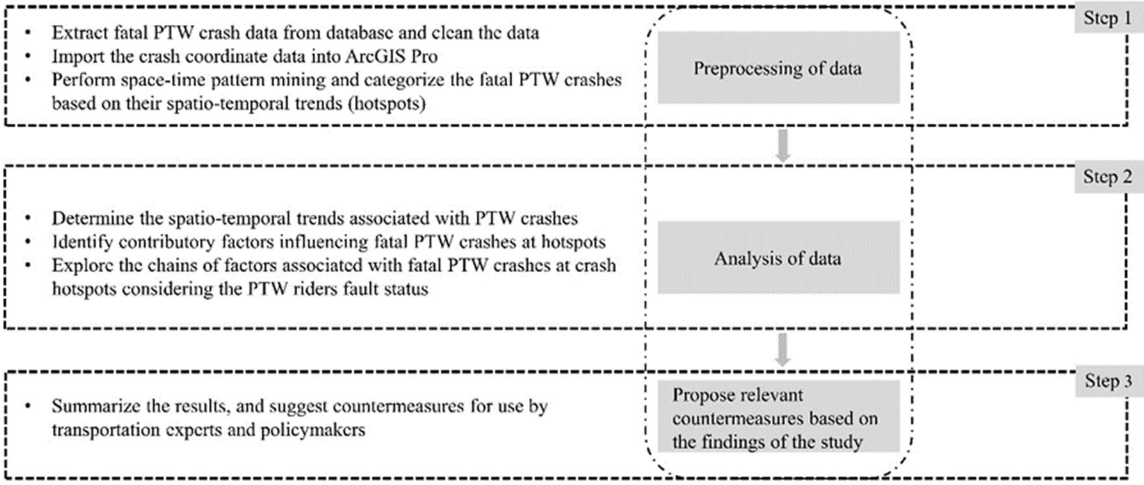

Figure 1. Proposed study framework.

2.2. Case Study Area and Data Description

2.2.1. Magnitude of the PTW Safety Problem in South Korea

The spatial scope of this study is South Korea, a country in Asia and home to about 51.7 million persons with a land area of 100,210 km^2 [39]. Its capital, Seoul, has a population density of 16,541 persons/km^2, extending across 605.2 km^2 [40]. The flexibility provided by PTWs and the incessant demand for quick food delivery services have led to their adoption by young people and delivery service operators in the country, particularly in Seoul, as one of the primary means of transport. In particular, the online food ordering service in South Korea grew from 2732 billion Korean Won in 2017 to 5273 billion Korean Won in 2018, and the number of PTWs per 1000 inhabitants in the country increased from 36.4 in 2012 to 40.0 in 2018 [41]. Research indicates that almost 60% of these PTWs are used for delivering food and parcels [42].

The increased use of PTWs has had a negative impact on traffic safety in South Korea. According to the WHO, there are about 4990 road-related deaths in South Korea each year, representing about 9.8 deaths/100,000 persons/year [43]. Regarding PTW crashes, it was estimated that there are about two deaths/100,000 persons/year, which is the fourth-highest in the same region and sixth-highest among similar high-income countries.

Studies focused on identifying the factors impacting the safety of PTWs in Korea are limited. The few studies in the literature highlight that overspeeding and non-compliance to general traffic rules are key factors influencing the severity of PTW crashes [44]. Although PTWs account for about 5% of traffic crashes, they are associated with a 12% fatality rate for all road crashes. The burden of PTW crashes in South Korea is unacceptably high and demands a conscientious and rigorous research approach to understand the factors influencing fatal PTW crashes for developing adequate measures to mitigate the impact of crashes.

2.2.2. Descriptive Statistics

In this study, police-recorded fatal PTW crashes on all rural and urban roads in South Korea between 2012 and 2017 were used. After data preprocessing, a total of 3555 fatal PTW crashes remained for analysis. It is worth noting that a fatal crash, as used in Korea for crash severity classification, is a crash in which death is recorded at the crash scene or up to 30 days after the crash [45]. Fatal PTW crashes were considered for this study due to their unacceptably high frequency and social cost.

The crash data contained detailed information pertaining to the temporal characteristics (year, season, time of day, and day of week), the crash characteristics (manner of collision, driver/rider violation, and crash location), the fault status of the driver/rider (at-fault and not-at-fault), the number of vehicles involved in the crash, the number of crash casualties, and the crash location coordinates (longitude and latitude). The crash data were geocoded and displayed on a map (Figure 2). It is noteworthy that other crash risk factors that may influence crashes, such as the drivers/riders' features, were not available in the crash dataset. All the data pertaining to the PTW crashes are summarized in Table 2.

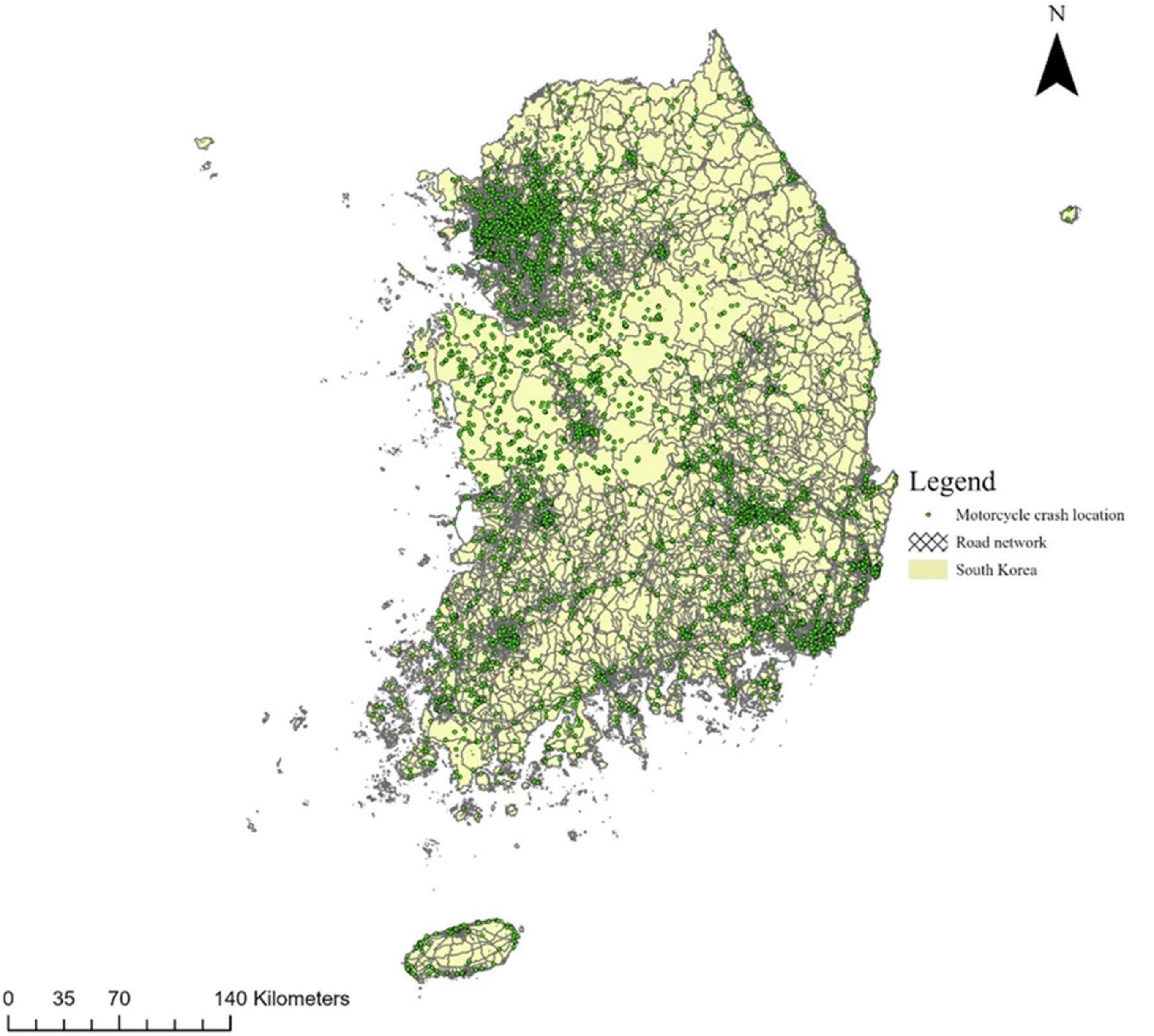

Figure 2. Visualization of the road network and PTW crash distribution in South Korea.

Figure 2 reveals that most of the crashes occurred around the capital city, Seoul. The table shows that the number of fatal PTW crashes is gradually increasing (2015–2017). In addition, fatal crashes are more prevalent during the summer and fall. Regarding the day-of-week variables, it was observed that a high percentage of crashes occurred on weekends. Most of the crashes were angle collisions and were caused by risky driving behaviors. In addition, most of the crashes occurred on main road sections, followed by intersections. PTW riders were primarily at fault for the crashes. As expected, more than half of the crashes resulted in multi-vehicle crashes with single-casualty outcomes.

Table 2. Variable description for fatal PTW crashes in South Korea (n = 3555 obs.).

Category	Frequency	Percentage	Category	Frequency	Percentage
Year			*Road segment (location)*		
2012	606	17.05	Bridge section	44	1.24
2013	564	15.86	At intersection	1128	31.73
2014	578	16.26	At pedestrian crossing	43	1.21
2015	570	16.03	Near pedestrian crossing	12	0.34
2016	616	17.33	Near intersection	332	9.34
2017	621	17.47	Main road	1994	56.09
Season			Other road segments (tunnel/underpass/overpass)	2	0.06
Fall (Sep.–Nov.)	1028	28.92	*At-fault party*		
Spring (Mar.–May)	930	26.16	Agricultural machinery	8	0.23
Summer (Jun.–Aug.)	1068	30.04	Bicycle	8	0.23
Winter (Dec.–Feb.)	529	14.88	Passenger car	651	18.31
Time of day			Construction machinery	39	1.10
Daytime (6 a.m.–5:59 p.m.)	2022	56.88	Engine-propelled bicycle	4	0.11
Nighttime (6 p.m.–5:59 a.m.)	1533	43.12	Freight truck	328	9.23
Day of week			Minibus/van	99	2.78
Monday	486	13.67	Other vehicle type	20	0.56
Tuesday	483	13.59	Two wheeler	2393	67.31
Wednesday	476	13.39	Unknown	5	0.14
Thursday	491	13.81	*Not-at-fault party*		
Friday	522	14.68	Agricultural machinery	8	0.23
Saturday	572	16.09	Bicycle	13	0.37
Sunday	525	14.77	Passenger car	597	16.79
Type of collision			Construction machinery	43	1.21
Angle	923	25.96	Engine-propelled bicycle	6	0.17
Crossing	102	2.87	Freight truck	326	9.17
Driving on edge of road	5	0.14	Minibus/van	138	3.88
Driving on road	14	0.39	No partner	1020	28.69
Head-on	351	9.87	Other vehicle type	17	0.48
Other crash types	726	20.42	Pedestrian	163	4.59
Rear-end	386	10.86	Two wheeler	1221	34.35
Rollover	393	11.05	Unknown	3	0.08
Run-off	108	3.04	*Number of vehicles involved*		
Sideswipe	233	6.55	One	1204	33.87
Collision on sidewalk	7	0.20	Two or more	2351	66.13
Collision at work zone	307	8.64	*Number of casualties*		
Violation			One	2786	78.37
Violation of intersection method	192	5.40	Two	595	16.74
Reckless driving/riding	2040	57.38	Three or more	174	4.89
Other violations	206	5.79			
Pedestrian protection violation	19	0.53			
Centerline crossing	358	10.07			
Safety distance violation	118	3.32			
Signal violation	546	15.36			
Over-speeding	76	2.14			

2.3. Methodology

This study seeks to identify chains of factor associations linked with at-fault and not-at-fault fatal PTW crashes at spatio-temporal crash hotspot locations. Unlike previous studies that manually cluster data for analysis or hotspot mining for analysis, this study identifies crash hotspots based on space and time dimensions, providing an enhanced understanding of how hotspots change over space and time. In addition, rather than applying traditional regression tools, which essentially identify the impact of individual factors on crash outcomes, this study applies a machine learning tool that identifies combinations of factors that collectively lead to a particular outcome. Using results from this method can provide more information for setting targeted countermeasures. This method is machine learning-based and requires no assumptions. The methods used for this research are described in the subsequent subsections.

2.3.1. Spatio-Temporal Pattern Mining

Space-time pattern mining tools available in ArcGIS were employed to explore the spatio-temporal trends in fatal PTW crashes. Fatal PTW crash data and other related spatial data are imported into ArcGIS. The data were then transformed into a space-time cube through a process of combining the crashes based on their spatial and temporal properties. To determine the spatial scope, a fishnet grid bin method was employed, where the height of the bins was determined based on a distance threshold in kilometers, which was established through a spatial autocorrelation test. For the temporal scope, a six-month time series was utilized, considering the available data and the size of the study area. The Mann–Kendall statistic was used to analyze the trends in the aggregated crash counts within each bin [46,47]. This resulted in each bin being associated with a z-score and a *p*-value that indicate the significance of the trend observed in the data.

The space-time cube is then inputted into the emerging hotspot analysis tool to discover the spatio-temporal hot and cold spots. The emerging hotspot tool can discover trends in the data by examining the space-time trends in the study area over a given period using the already-created space-time cube file. The researcher must select appropriate neighborhood distance and time step parameters, which would be used to determine how many surrounding bins are examined together spatio-temporally when assessing the local space-time clustering. They are subsequently employed to estimate the Getis-Ord G_i^* statistic for each bin. This statistic is specified in Equation (1) below.

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij} x_j - \bar{X} \sum_{j=1}^{n} w_{ij}}{S\sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{ij}\right)^2\right]}{n-1}}}$$
$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$
$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \quad (1)$$

where x_j denotes the bin value for feature j, the spatial weight between features i, and j is w_{ij}; n shows the number of bins, $\bar{X}$ represents the mean of the attributes, and S is the corresponding standard deviation.

Upon running the tool, each bin is associated with a hotspot classification, a z-score, and a *p*-value which explains the trend in each location. The emerging hotspot analysis tool then categorizes the study area using each crash location's resultant trends' z-score and *p*-value, and each bin's hotspot z-score and *p*-value. The distance interval and time step interval used in building the space-time cube are one km and six months, respectively. This led to the creation of 168,768 bins for analysis (minimum time-step interval—10; maximum number of bins—2 billion), with a resulting 0.55% temporal bias. In addition, the emerging hotspot analysis used a 15 km neighborhood distance and three neighborhood time steps.

This approach classified locations within the study area into two main groups: no pattern detected (analysis location is neither hot nor cold) and critical hotspot (analysis location is a statistically significant hotspot needing urgent attention). The data associated with the group of interest (critical hotspots) were extracted for further exploration.

2.3.2. Association Rule Mining

This study employed the Association Rules Mining (ARM) technique, a data mining approach first proposed by Agrawal et al. to discover frequent chains of items that occur together in an event. It has since gained popularity among transportation researchers due to its ease of understanding and practicality in terms of efficiently discovering the relationships between different crash risk factors in a database [48]. In addition, since the algorithm employs the relative frequency approach in identifying frequent itemsets (crash risk factors) and combinations with each other, it is more suited for unbalanced data such as crash data.

Researchers have highlighted that the ARM technique is particularly advantageous since it requires no predefined underlying assumptions, has the capability of handling small datasets, and requires no specification of dependent and independent variables before analysis. In particular, previous transport safety studies employed ARM to explore small datasets ranging from 126 to 449 crash observations [49].

Suppose the fatal PTW crash database is $D = \{d_1, d_2, \ldots, d_y\}$, where each crash observation consists of a subset of x unique items/variables contained in the itemset I, such that, $I = \{i_1, i_2, \ldots, i_x\}$. Rules generated by the ARM technique can be defined as an implication of the form $P => Q$, such that $P, Q \subseteq I$, and $P \cap Q = \emptyset$. The left-hand side (LHS) of the rule, P, is known as the antecedent, and the right-hand side of the rule, Q, is the consequent. For example, the rule $P => Q$ can be explained as follows: if P exists as a crash risk factor pertaining to a certain crash observation, then Q is likely to occur.

Obtaining useful rules involves a process guided by three main measures: support, confidence, and lift [48]. Support of a rule shows how popular the itemset is by expressing the proportion of the entire database covered by the rule. Support of a rule is expressed as follows:

$$Support(P => Q) = \frac{|P \cap Q|}{|D|} \tag{2}$$

where $|\mathrm{P} \cap \mathrm{Q}|$ shows the number of times itemsets P and Q co-occur, and $|D|$ represents the number of crashes in database D. Confidence of a rule depicts the percentage of cases where the consequent of the rule, Q, occurs given that the antecedent, P, already occurred. Confidence of the rule $P => Q$ is estimated as follows [50]:

$$Confidence(P => Q) = \frac{Support(P => Q)}{Support(P)} \tag{3}$$

The support–confidence framework is complemented by the measure, lift, to efficiently prune infrequent itemsets from the dataset [51]. The lift of a rule $P => Q$ shows the how frequently the antecedent and consequent simultaneously occur. It can also be described as how likely an itemset Q occurs given that P has occurred, while controlling for the popularity of itemset Q (serving as a measure of the statistical interdependence of a rule) [15]. It is computed as follows:

$$Lift(P => Q) = \frac{Support(P => Q)}{Support(P) \cdot Support(Q)} \tag{4}$$

A rule is deemed interesting and valuable when the lift is greater than 1 (indicating positive interdependence between both P and Q). In addition, there is no correlation between both P and Q when the lift is 1. A negative interdependence is established when the lift of the rule is less than 1. It is noteworthy that higher lift values connote higher interestingness of the rule.

In ARM, no specific rules govern the selection of support, confidence, and lift values. However, it is desirable to have rules with sufficiently high support, large confidence, and a lift value higher than one. Researchers have advocated for the validity verification of the addition of each item i to the rule to ensure that the addition of a new item results in an increase in the lift [15]. To do this, a rule with item i in the antecedent itemset is used as the starting point. The lift value, $Lift_{A_{i+1}}$, obtained after adding a new item $i+1$ is then compared to the previous lift value, $Lift_{A_i}$, as follows:

$$LIC = \frac{Lift_{A_i}}{Lift_{A_{i+1}}} \quad (5)$$

where A_i is the antecedent with i item, and A_{i+1} is the antecedent after a new item is added. A rule with LIC greater than the preselected minimum threshold is more favored.

There are no specific rules governing the selection of support (s), confidence (c), lift (l), and LIC threshold values; however, consistent with transport safety literature and in view of the data used for the analysis, the following thresholds are used: $s = 1\%$, $c = 5\%$, $l = 1.10$, and $LIC = 1.02$. The rule mining was conducted using R software [52].

3. Results and Discussions

3.1. Spatio-Temporal Pattern Mining

Spatio-temporal trends in fatal PTW crashes were discovered to help identify recent trends and where fatal PTW crashes are frequently clustered. The spatio-temporal patterns of fatal PTW crashes occurring on urban and rural roads in South Korea between 2012 and 2017, as illustrated in Figure 3, show that fatal crashes are concentrated around the capital, Seoul, and its environs. Thus, the analysis results would focus on those areas.

From Figure 3, it is evident that consecutive hotspots (indicated in light brown color) are predominant in Seoul, particularly around the central business district. Consecutive hotspots are those locations in which statistically significant hotspot bins occur uninterruptedly in recent time steps. Still, within Seoul, it is observed that few locations are sporadic fatal PTW hotspots. In essence, these locations are hot in some years and not hot in other years. Essentially, none of the time steps in these bins have ever been statistically significant cold spots. Note that the yellow-colored sections are those with no crashes, and those white-colored sections are those where crashes occurred, but no patterns were identified, likely due to the low frequency of crashes (these are dotted across the entire map and clearly visible when zoomed in to, as shown below).

The neighboring cities around Seoul, namely Incheon, Gyeonggi, Sejong, and Gimpo, also had some hotspot locations. While Incheon, Gyeonggi, and Sejong had few consecutive hotspot locations, Gimpo had new and emerging hotspots. New hotspots are identified by bins that have never been hot but had their final time steps as statistically significant hotspots. The other cities had no statistically significant hotspots or cold spots, indicating that Seoul and its environs are critical locations worth studying. The findings from the spatio-temporal hotspot analysis are intuitive as the patronization of food delivery service using PTWs in Seoul, South Korea's most densely populated city and home to many companies, is the order of the day as most Koreans like to eat away from home [53]. These delivery services are recently gaining popularity in the neighboring cities as population flows to cities around Seoul are being observed due to the government's promotion of decentralization and population redistribution [39].

Table 3 shows the number of crashes classified at each location and the number of crashes per classification category with statistically positive or negative trends determined based on their z-scores and p-values. Crashes in locations with significant and positive trends have positive z-scores with a p-value < 0.1. Interestingly, almost all hotspots showed positive/increasing trends, indicating the need for more focus on these areas and suggesting interventions for use by road safety authorities.

Out of the 3555 crashes considered, 198 crashes were identified at fatal PTW crash hotspots. In particular, 181 crash observations were identified at consecutive hotspots, and only one crash observation found at a new hotspot showed statistically positive trends (182 crashes in total with positive trends at hotspots). Since most crashes were spread across the entire country, they were identified as no-pattern locations. As this study seeks to identify the key chains of factors associated with PTW crashes at hotspot locations, the 182 crash observations at the significant hotspot locations were extracted for further analysis. A total of 128 fatal crashes were caused by the PTW riders, and the remaining were not caused by the PTW riders. The dataset used for the analysis is small. However, the machine learning method is robust at identifying patterns in small datasets—an advantage over the traditional regression methods [54].

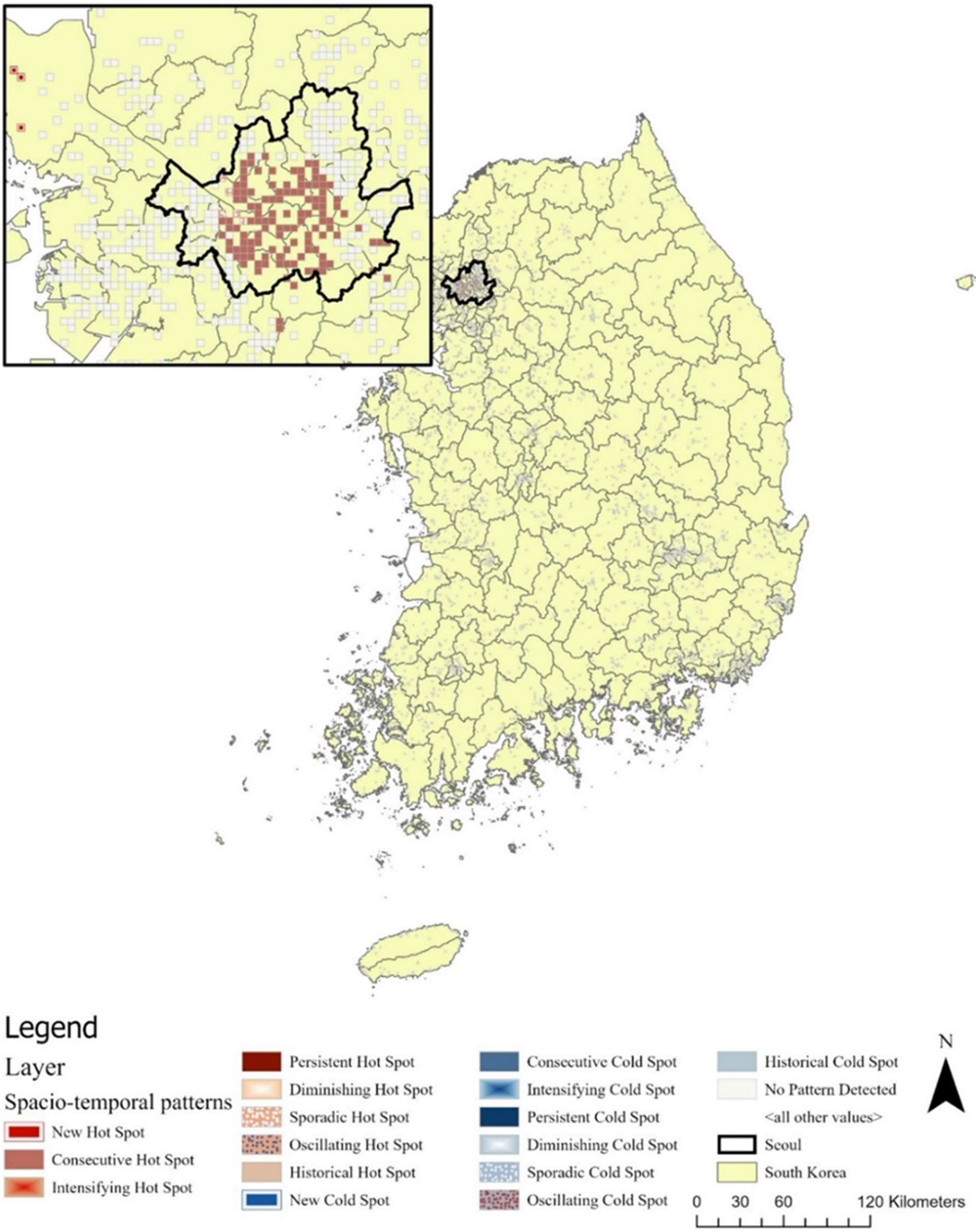

Figure 3. Spatio-temporal hotspots of fatal PTW crashes in South Korea (2012–2017).

Table 3. Space-time trend mining results.

Emerging Hotspot Classification	Negative Trend		Positive Trend		Total (Significant Trends)	Total (Non-Significant Trends)	Grand Total
	Significant Trends at 90% CL	Non-Significant Trends at 90% CL	Significant at 90% CL	Non-Significant at 90% CL			
Consecutive Hotspot	0	0	181	13	181	13	194
New Hotspot	0	1	1	3	1	4	5
No Pattern Detected	589	1626	773	360	1362	1986	3348
Sporadic Hotspot	0	0	0	8	0	8	8
Grand Total	589	1627	955	384	1544	2011	3555

S = 0.01; C = 0.05; L = 1.1, LIC = 1.02.

3.1.1. Interesting Rules for All Fatal PTW Crashes at Critical Locations

The association rules for fatal PTW crashes at hotspots are shown in Table 4. All the rules satisfy the minimum lift and LIC thresholds, signifying greater interdependence between the antecedents and consequents of the rules. The first rule in Table 4 shows that fatal PTW crashes caused by the fault of freight truck drivers are likely to occur on Mondays. Essentially, a freight truck driver being at fault increases the probability that a PTW crash will occur on a Monday by 3.92 times. According to rule 2, if the truck driver disregarded the safety distance regulation, the likelihood of a PTW crash increases 6.54 times more than the average. Freight transport share by road in Korea increased by more than 90% since 2011, and the number of trucks on the roads also keeps increasing. Since truck operation times in Korea is during the daytime (9 a.m.–6 p.m.), the likelihood of interactions with PTWs on the road during the daytime is likely to be high, particularly on Mondays, when many trucks and PTW delivery workers operate. In addition, in line with the literature, the likelihood of crashes in Korea (particularly involving trucks) has been identified to increase with safe distance violations [55], which is likely due to aggressive driving, inattention, or high traffic volume. These driving behaviors are likely to impact PTW safety negatively.

Table 4. List of interesting rules identified for fatal PTW at critical locations.

No.	Rules	S%	C%	Lift	LIC
1	{At-fault party: freight truck} => {Day of week: Monday}	1.64	60.00	3.92	-
2	{At-fault party: freight truck, Violation: safety distance violation} => {Day of week: Monday}	1.09	100.00	6.54	1.67
3	{Road segment: near intersection} => {Violation: centerline crossing}	3.83	41.18	3.01	-
4	{Road segment: near intersection, Type of collision: head-on} => {Violation: centerline crossing}	1.09	66.67	4.88	1.62
5	{Road segment: near intersection, Year: 2012} => {Violation: centerline crossing}	1.64	100.00	7.32	2.43
6	{Road segment: near pedestrian crossing} => {Type of collision: angle}	1.64	75.00	2.98	-
7	{Road segment: near pedestrian crossing, Violation: centerline crossing} => {Type of collision: angle}	1.09	100.00	3.98	1.33
8	{Not-at-fault party: minibus/van} => {Type of collision: angle}	3.28	66.67	2.65	-
9	{Not-at-fault party: minibus/van, Violation: signal violation} => {Type of collision: angle}	1.64	100.00	3.98	1.50
10	{Road segment: at pedestrian crossing} => {Violation: signal violation}	1.09	50.00	2.41	-
11	{Road segment: at pedestrian crossing, Year: 2013} => {Violation: signal violation}	1.09	100.00	4.82	2.00
12	{Road segment: bridge section} => {Type of collision: other crash type}	2.19	66.67	2.39	-
13	{Road segment: bridge section, Year: 2017} => {Type of collision: other crash type}	1.09	100.00	3.59	1.50
14	{Violation: over-speeding} => {Season: summer}	2.73	71.43	2.33	-
15	{Violation: over-speeding, Year: 2013} => {Season: summer}	1.09	100.00	3.27	1.40
16	{At-fault party: freight truck} => {Not-at-fault party: PTW}	2.19	80.00	2.32	-
17	{At-fault party: freight truck, Violation: safety distance violation} => {Not-at-fault party: PTW}	1.09	100.00	2.90	1.25
18	{Violation: centerline crossing} => {Not-at-fault party: passenger car}	5.46	40.00	2.15	-
19	{Violation: centerline crossing, Year: 2012} => {Not-at-fault party: passenger car}	2.73	83.33	4.49	2.08
20	{Violation: centerline crossing, Year: 2014} => {Not-at-fault party: passenger car}	1.09	50.00	2.69	1.25
21	{Type of collision: rear-end} => {At-fault party: passenger car}	3.83	46.67	2.14	-
22	{Type of collision: rear-end, Violation: over-speeding} => {At-fault party: passenger car}	1.09	100.00	4.58	2.14
23	{Type of collision: rear-end, Year: 2013} => {At-fault party: passenger car}	2.19	66.67	3.05	1.43
24	{Violation: signal violation} => {Type of collision: angle}	10.38	50.00	1.99	-
25	{Violation: signal violation, Year: 2012} => {Type of collision: angle}	2.19	100.00	3.98	2.00
26	{Violation: signal violation, Year: 2015} => {Type of collision: angle}	2.19	80.00	3.18	1.60
27	{Type of collision: angle} => {Road segment: at intersection}	13.66	54.35	1.95	-
28	{Type of collision: angle, Violation: signal violation} => {Road segment: at intersection}	9.84	94.74	3.40	1.74

S = 0.01; C = 0.05; L = 1.1, LIC = 1.02.

Rule 3 depicts that if a PTW crash occurs near an intersection, it is likely that it was due to a centerline crossing violation. According to rule 4, if the PTW crash occurs near an intersection and it is a head-on collision, then the probability of it being caused by a centerline violation is 4.88 times more than the average. In addition, these kinds of crashes mainly occurred in 2012 (Rule 5). These results are plausible as violating centerline regulations, particularly at intersection areas where interactions between different road user types are high, could lead to head-on collisions with unsuspecting drivers/riders.

Other rules in Table 4 present interesting findings. PTW crashes involving not-at-fault minibus/van drivers will likely be an angle collision (lift = 2.65). In addition, when a not-at-fault minibus/van crash occurs due to the PTW rider disregarding the traffic signal, the likelihood of angle collisions is 3.98 times more than the average (rule 9). From rule 14, it is observed that over-speeding violations increase the probability of PTW crashes in summer by 2.33 times, which is supported by the literature and is intuitive as drivers, in general, are likely to engage in risky behaviors when they perceive the weather to be good [49]. Rules 18–20 show that if a PTW-passenger car crash occurs due to a centerline violation, the PTW rider is likely to be at fault (not the passenger car driver). However, according to rule 21, a passenger car driver is likely to be at fault in rear-end crashes (lift = 2.14). From rule 22, if the rear-end collision is caused by over-speeding, the probability that a passenger car driver is at fault increases 4.58 times more than the average. A survey showed that drivers are likely to step on their brakes about 70% of the time when they feel threatened. An unsuspecting PTW rider who happens to be right behind this vehicle could easily run into it if the car driver slams his brakes and stops abruptly [56].

Finally, from rule 24, signal violations are likely to lead to angle collisions (supporting rule 9—{Not-at-fault party: minibus/van, Violation: signal violation} => {Type of collision: angle}). In addition, Rules 27 and 28 suggest that angle collisions occurring due to signal violations are likely to occur at intersection segments. This finding is in line with the literature. Research shows that running red lights at intersections increases the vulnerability of PTW riders and raises the chance of angle collisions [57]. The authors added that the vulnerability of PTW riders to angle collisions could be substantially reduced by introducing red-light cameras.

3.1.2. Interesting Rules for Fatal PTW Crashes Based on PTW Rider's Fault Status

To obtain a clearer understanding of the risk factors influencing fatal PTW crashes at crash hotspots due to the fault of PTW riders, association rules were mined for PTW rider-at-fault crashes occurring at hotspot locations (see Table 5). Rule 1 shows that PTW crashes occurring on Saturdays are likely to be the fault of PTW riders (lift = 1.18). According to rules 2 and 3, crashes occurring on Saturdays at main road segments or due to reckless riding are likely due to the fault of PTW riders. In particular, PTW crashes in Korea mainly occurred on Saturdays [58]. Due to the reduced traffic volumes, PTW riders are likely to ride recklessly, causing crashes.

The crash risk factors contained in rules 4–7 show that PTW crashes resulting in two casualties due to reckless riding during the nighttime or at other road segments, such as tunnels, overpasses, or underpasses, are likely to be due to the fault of PTW riders. According to rule 8, if a PTW crash occurs at these segments, the chance that the PTW rider is at fault increases 1.16 times. In addition, when the crash occurs at other road segments due to reckless riding, during the nighttime or summer, the probability that the PTW rider is at fault increases by 1.28, 1.34, and 1.25 times, respectively (see rules 9, 10, 11, and 12). Finally, the risk factors in rules 13 to 18 show strong associations between nighttime crashes, reckless riding, and PTW rider at-fault variables. In South Korea, most PTW crashes occur during the nighttime from 6 p.m. to 6 a.m. [58]. Most ride fast and weave through traffic to reach their destinations on time. Previous research identified riding at night and on weekends as a predictor of offenses among PTW riders [59]. These offenses are likely to increase the risk of crashes involving PTWs, putting the riders and their passengers at a greater risk of fatal crashes. Interestingly, a study conducted in Singapore maintained that PTW riders are likely not to be at fault for a crash that occurs on the expressway or at intersections during the nighttime due to the low conspicuity of the rider [17]. However, in South Korea, PTWs are banned from using expressways. Thus, this study considered all rural and urban roads aside from expressways. Additionally, note that this study focused on crashes at hotspot locations, which may be the reason for the difference in findings.

Table 5. List of interesting rules identified for fatal PTW rider at-fault crashes at critical locations.

No.	Rules	S%	C%	Lift	LIC
1	{Day of week: Saturday} => {At-fault party: PTW}	13.11	82.76	1.18	-
2	{Day of week: Saturday, Road segment: other road segments} => {At-fault party: PTW}	8.20	93.75	1.34	1.13
3	{Day of week: Saturday, Violation: reckless driving/riding} => {At-fault party: PTW}	9.29	94.44	1.35	1.14
4	{Number of casualties: two} => {At-fault party: PTW}	18.58	80.95	1.16	-
5	{Number of casualties: two, Road segment: other road segments} => {At-fault party: PTW}	10.93	86.96	1.24	1.07
6	{Number of casualties: two, Time of day: nighttime} => {At-fault party: PTW}	14.75	87.10	1.25	1.08
7	{Number of casualties: two, Violation: reckless driving/riding} => {At-fault party: PTW}	8.20	93.75	1.34	1.16
8	{Road segment: other road segments} => {At-fault party: PTW}	43.72	80.81	1.16	-
9	{Road segment: other road segments, Violation: reckless driving/riding} => {At-fault party: PTW}	31.69	89.23	1.28	1.10
10	{Road segment: other road segments, Time of day: nighttime} => {At-fault party: PTW}	33.88	87.32	1.25	1.08
11	{Road segment: other road segments, Time of day: nighttime, Violation: reckless driving/riding} => {At-fault party: PTW}	24.04	93.62	1.34	1.07
12	{Road segment: other road segments, Season: summer} => {At-fault party: PTW}	14.75	87.10	1.25	1.08
13	{Time of day: nighttime} => {At-fault party: PTW}	53.01	76.98	1.10	-
14	{Time of day: nighttime, Violation: reckless driving/riding} => {At-fault party: PTW}	31.69	89.23	1.28	1.16
15	{Time of day: nighttime, Year: 2015} => {At-fault party: PTW}	9.84	85.71	1.23	1.11
16	{Time of day: nighttime, Type of collision: other crash type} => {At-fault party: PTW}	16.94	81.58	1.17	1.06
17	{Time of day: nighttime, Type of collision: other crash type, Violation: reckless driving/riding} => {At-fault party: PTW}	12.57	88.46	1.26	1.08
18	{Time of day: nighttime, Year: 2013} => {At-fault party: PTW}	12.57	79.31	1.13	1.03

S = 0.01; C = 0.05; L = 1.1, LIC = 1.02.

Association rules for PTW crashes that were not caused by the PTW rider's fault were also explored (see Table 6). This was performed by setting the item {Not-at-fault party: PTW} as the consequence of the rule and identifying strong rules that meet the lift and LIC criteria. Due to the low number of crashes in which the PTW rider was not at fault for the crash, the support of the rules generated was lower compared to the other sets of rules generated. The crash risk factors in rules 1 to 5 show that PTW riders are likely not to be at fault for rear-end and head-on fatal crashes at hotspot locations. This result seems counterintuitive as research conducted in Iran identified that the at-fault probability of PTW riders increases in head-on and rear-end crashes [32]. However, from rule 6, it is evident that the probability that the PTW rider is not at fault increases by 1.94 times if the head-on crash is caused by the other drivers' signal violation/red-light running act. This finding is plausible as the number of red-light running crashes involving cars in Korea is substantial. A study conducted in Cheongju, Korea, identified that out of 6130 crashes, 2246 were red-light running crashes, of which private cars formed the majority (58%) [60]. According to the authors, the reasons why drivers run red lights could be drowsiness, inattentiveness, or a perception of being delayed while waiting. As a result, an unsuspecting rider is likely to collide with a red-light-running vehicle, particularly at an intersection. In line with our study, Campbell et al. identified that disregarding traffic signals led to more head-on crashes than other crash types [61].

According to rule 7, a PTW rider is likely not to be at fault if two or more vehicles are involved in the crash (lift = 1.50). Items such as daytime, sideswipes, spring, and safety distance violation are common among the items in rules 8 to 18. This finding shows that, at critical hotspot locations, the PTW rider is likely not to be at fault if s/he is involved in two or more vehicle crashes, particularly sideswipes, during the daytime or spring due to improper safety distance. This finding is likely, as several interactions occur between different road users during the daytime. Drivers who do not look carefully before changing lanes are likely to obstruct or hit an oncoming PTW. These findings are consistent with other studies. For example, the literature identified that a PTW rider is likely not at fault if a sideswipe crash occurs during the day in an urban setting [32].

Table 6. List of interesting rules identified for fatal PTW rider not-at-fault crashes at critical locations.

No.	Rules	S%	C%	Lift	LIC
1	{Type of collision: rear-end} => {Not-at-fault party: PTW}	4.92	60.00	1.74	-
2	{Type of collision: rear-end, Year: 2013} => {Not-at-fault party: PTW}	2.73	83.33	2.42	1.39
3	{Type of collision: head-on} => {Not-at-fault party: PTW}	5.46	55.56	1.61	-
4	{Type of collision: head-on, Year: 2017} => {Not-at-fault party: PTW}	1.64	75.00	2.18	1.35
5	{Type of collision: head-on, Year: 2016} => {Not-at-fault party: PTW}	2.19	66.67	1.94	1.20
6	{Type of collision: head-on, Violation: signal violation} => {Not-at-fault party: PTW}	2.19	66.67	1.94	1.20
7	{Number of vehicles involved: two or more} => {Not-at-fault party: PTW}	34.43	51.64	1.50	-
8	{Number of vehicles involved: two or more, Violation: safety distance violation} => {Not-at-fault party: PTW}	5.46	71.43	2.07	1.38
9	{Number of vehicles involved: two or more, Road segment: near pedestrian crossing} => {Not-at-fault party: PTW}	1.09	66.67	1.94	1.29
10	{Number of vehicles involved: two or more, Road segment: bridge section} => {Not-at-fault party: PTW}	2.19	66.67	1.94	1.29
11	{Number of vehicles involved: two or more, Type of collision: sideswipe} => {Not-at-fault party: PTW}	4.37	66.67	1.94	1.29
12	{Number of vehicles involved: two or more, Time of day: daytime} => {Not-at-fault party: PTW}	14.75	65.85	1.91	1.28
13	{Number of vehicles involved: two or more, Time of day: daytime, Violation: safety distance violation} => {Not-at-fault party: PTW}	3.28	85.71	2.49	1.30
14	{Number of vehicles involved: two or more, Time of day: daytime, Violation: reckless driving/riding} => {Not-at-fault party: PTW}	4.92	81.82	2.38	1.24
15	{Number of vehicles involved: two or more, Time of day: daytime, Year: 2013} => {Not-at-fault party: PTW}	2.19	80.00	2.32	1.21
16	{Number of vehicles involved: two or more, Season: spring, Violation: safety distance violation} => {Not-at-fault party: PTW}	1.64	75.00	2.18	1.24
17	{Number of vehicles involved: two or more, Season: spring, Year: 2015} => {Not-at-fault party: PTW}	1.64	75.00	2.18	1.24
18	{Number of vehicles involved: two or more, Season: spring, Time of day: daytime} => {Not-at-fault party: PTW}	4.37	72.73	2.11	1.20

S = 0.01; C = 0.05; L = 1.1, LIC = 1.02.

4. Conclusions

4.1. Key Findings

This study's ultimate goal was to investigate the associations between chains of factors that lead to fatal PTW crashes at crash hotspot locations in South Korea, considering the fault status of the PTW rider. Although previous studies highlighted that the influence of PTW crash risk factors could differ based on the fault status of the rider and that crash hotspots are critical locations with particularly high crash frequencies, no study comprehensively examined the key associations between crash risk factors influencing fatal PTW crashes at hotspot locations considering the fault status of the rider and the spatio-temporal locations of the crash. This study bridged the literature gap by

- Identifying different types of fatal PTW crash hotspots in South Korea based on spatio-temporal information, together with their evolution patterns over time;
- Exploring the chains of contributory factors of fatal PTW crashes at the crash hotspots;
- Discovering and comparing the chains of contributory factors influencing fatal PTW crashes considering the fault status of the PTW rider.

To achieve this study's objective, the police-reported fatal PTW crashes were obtained, and time-stamped crash observations were geocoded and displayed on a map. This information is then used in spatio-temporal clustering of fatal PTW crashes to identify PTW crashes at PTW crash hotspots with a significantly increasing trend. The association rules mining technique was then employed to explore the associations existing among the crash risk factors and between crash risk factors and the at-fault involvement of the PTW rider.

Generally, this study identified that the majority of the significant fatal PTW hotspot locations were in the Seoul Metropolitan Area. While the hotspot locations in Seoul were predominantly consecutive, those in Gimpo were mostly new hotspots. From the general association rules, it was identified that signal violations are likely to cause angle collisions and have strong associations with intersection crashes. In addition, violations such as running red lights and centerline crossing are strongly related to angle collisions. Over-speeding is also found to have strong associations with PTW crashes during the summer, and rear-end PTW-involved crashes occurring due to over-speeding are likely to be caused by passenger car drivers.

Upon analyzing the rules identified at hotspots for both at-fault and not-at-fault PTW riders, it was identified that while there were strong associations between fatal PTW rider at-fault crashes and the nighttime variable at crash hotspots, that of PTW rider not-at-fault crashes were highly correlated with the daytime variable. Regarding traffic violations leading to fatal PTW crashes at hotspot locations, while reckless riding is the predominant and the only violation found among the rules mined for PTW rider at-fault crashes, those for PTW rider not-at-fault crashes were diverse, including signal violations, safety distance violations, and reckless driving. Concerning the season in which fatal PTW crashes occurred, it was discovered that summer had strong associations with PTW rider at-fault crashes, and spring had strong correlations with PTW rider not-at-fault crashes.

4.2. Proposed Countermeasures

Based on the findings of this study, the following countermeasures are proposed for both PTW riders and other drivers to help mitigate the fatality of PTW crashes at the hotspot locations.

- The finding that most fatal PTW crash locations in Seoul have begun getting more popular in recent years and the significant positive trends identified highlight that the safety of PTW riders is gradually worsening. Therefore, there is a need to intensify rider/driver education and enforcement to help check traffic violations such as speeding, signal violations, and centerline violations to help reduce the fatality of these crashes. Education of road users could be undertaken through safety campaigns, and enforcement could be targeted at both the road users and the business owners who use PTWs for business;
- Excessive speeds by both riders and drivers could be managed by the application of traffic calming designs, such as introducing mini traffic circle configurations at intersections, speed humps, roadway diets, and speed limit reductions, where necessary, on rural and urban roads. In addition, other control measures which self-enforce speed reductions, such as horizontal and vertical deflections, could be explored to help reduce the approach speed of vehicles;
- While signal violations could be checked by deploying adequate red-light cameras and enforcement measures, it would be worthwhile to consider making traffic signals more visible for all road users. This can be achieved by carefully considering their design, placement, and luminance when choosing traffic signals. Clearly visible roadway signs should be placed at vantage points to remind riders of impending intersections and speed limits. In addition, adequately putting in measures to control the speeds of vehicles could reduce their chance of running red lights;
- Centerline violations at crash hotspots could be controlled by using countermeasures such as centerline rumble strips to alert drivers who may unintentionally drive across the centerlines. In addition, illuminated in-ground light-emitting diodes (LEDs) and retroreflective lane markings could be used in marking out the centerlines to help provide visual guidance and increase drivers' awareness of road separations;

- The finding that fatal rider at-fault PTW crashes at hotspot locations is strongly associated with nighttime highlights that these riders may be fatigued, sleepy, or under the influence of drugs due to the need to overwork and achieve daily targets. Again, this problem could be tackled by educating riders on the need to rest when tired. While rumble strips and better lighting at night could be used to check crashes associated with drowsiness and low visibility, nighttime riding restrictions at hotspots could be meted out to PTW riders, particularly delivery company riders who are found culpable to ensure that riders obey traffic regulations even at night. In addition, there is a need to encourage the wearing of luminous vests to increase the conspicuity of riders. Similar interventions could be explored in checking PTW rider at-fault crashes during summer;
- Other vehicle drivers have been identified to have strong associations with traffic infringements such as reckless driving, signal violations, and safety distance violations, leading to PTW rider not-at-fault crashes. These violations are likely due to driver distraction and aggressive driving. To control rider not-at-fault crashes, vehicle drivers need to be trained on the need to stay focused and be on the lookout for PTW riders, especially when making lane changes. Additionally, there is a need to heighten enforcement such that repeat offenders of traffic infringements are subjected to harsher punishments to deter others;
- In the long term, interventions such as providing exclusive PTW lanes at crash hotspots and creating policies for anti-lock braking systems should be considered.

4.3. Study Limitations and Future Study Areas

The findings obtained from this comprehensive study provide a clearer insight into the factors impacting fatal PTW crashes at hotspots considering PTW riders' fault status. Together with the recommended countermeasures, this study can be useful for policymakers and road safety authorities to develop appropriate strategies geared toward reducing fatal PTW crashes at hotspots. As a limitation, this study lacked basic information about the driver/rider and other environmental features. These variables should be included in future studies to help understand how they affect the occurrence of PTW crashes clearly at hotspot locations. In addition, selecting the most relevant rules for a given application can be challenging, as it is largely based on the researcher's knowledge of the field and the thresholds selected. Although this study used a dataset from South Korea, its findings and recommendations could be extended to other countries with increasing dependence on PTWs. Future studies may adopt the method adopted in this study to conduct a more comprehensive analysis to understand and provide more targeted mitigation strategies specific to their locations. It is also recommended that future studies consider exploring the factors associated with other crash severity types at spatio-temporal hotspot locations, considering variables like land use. Finally, given that Seoul is a megacity with the highest population density in the country, it would be worthwhile to conduct a separate analysis to identify hotspots in Seoul and compare them to other parts of the country.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. WHO (World Health Organization). Death on the Roads. Available online: https://extranet.who.int/roadsafety/death-on-the-roads//#country_or_area/KOR (accessed on 26 February 2023).
2. Al-Ahmadi, H.M.; Jamal, A.; Ahmed, T.; Rahman, M.T.; Reza, I.; Farooq, D. Calibrating the Highway Safety Manual Predictive Models for Multilane Rural Highway Segments in Saudi Arabia. *Arab. J. Sci. Eng.* **2021**, *46*, 11471–11485. [CrossRef]

3. WHO. *Powered Two- and Three-Wheeler Safety: A Road Safety Manual for Decision-Makers and Practitioners*; WHO: Geneva, Switzerland, 2017.
4. WHO. *Global Status Report on Road Safety 2018*; WHO: Geneva, Switzerland, 2018.
5. NHTSA. *Traffic Safety Facts 2005: Motorcycles*; National Highway Traffic Safety Administration: Washington, DC, USA, 2007.
6. Johnston, P.; Brooks, C.; Savage, H. *Fatal and Serious Road Crashes Involving Motorcyclists*; Department of Infrastructure, Transport, Regional Development and Local Government: Canberra, Austrilia, 2008.
7. OECD/ITF. *Improving Safety for Motorcycle, Scooter and Moped Riders*; ITF Research Reports; OECD: Paris, France, 2015; ISBN 9789282107935.
8. Tiwari, G. Safety challenges of powered two-wheelers. *Int. J. Inj. Control Saf. Promot.* **2015**, *22*, 281–283. [CrossRef] [PubMed]
9. Aidoo, E.N.; Amoh-Gyimah, R. Modelling the risk factors for injury severity in motorcycle users in Ghana. *J. Public Health* **2020**, *28*, 199–209. [CrossRef]
10. Murphy, P.; Morris, A. Quantifying accident risk and severity due to speed from the reaction point to the critical conflict in fatal motorcycle accidents. *Accid. Anal. Prev.* **2020**, *141*, 105548. [CrossRef]
11. Talbot, R.; Brown, L.; Morris, A. Why are powered two wheeler riders still fatally injured in road junction crashes?—A causation analysis. *J. Saf. Res.* **2020**, *75*, 196–204. [CrossRef] [PubMed]
12. Theofilatos, A.; Yannis, G.; Antoniou, C.; Chaziris, A.; Sermpis, D. Time series and support vector machines to predict powered-two-wheeler accident risk and accident type propensity: A combined approach. *J. Transp. Saf. Secur.* **2018**, *10*, 471–490. [CrossRef]
13. De Rome, L.; Senserrick, T. Factors Associated with Motorcycle Crashes in New South Wales, Australia, 2004 to 2008. *Transp. Res. Rec. J. Transp. Res. Board* **2011**, *2265*, 54–61. [CrossRef]
14. Abdul Manan, M.M.; Várhelyi, A.; Çelik, A.K.; Hashim, H.H. Road characteristics and environment factors associated with motorcycle fatal crashes in Malaysia. *IATSS Res.* **2018**, *42*, 207–220. [CrossRef]
15. Montella, A.; de Oña, R.; Mauriello, F.; Rella Riccardi, M.; Silvestro, G. A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. *Accid. Anal. Prev.* **2020**, *134*, 105251. [CrossRef]
16. Al Hamami, M.; Matisziw, T.C. Measuring the spatiotemporal evolution of accident hot spots. *Accid. Anal. Prev.* **2021**, *157*, 106133. [CrossRef]
17. Haque, M.M.; Chin, H.C.; Huang, H. Modeling fault among motorcyclists involved in crashes. *Accid. Anal. Prev.* **2009**, *41*, 327–335. [CrossRef] [PubMed]
18. Bjørnskau, T.; Nævestad, T.-O.; Akhtar, J. Traffic safety among motorcyclists in Norway: A study of subgroups and risk factors. *Accid. Anal. Prev.* **2012**, *49*, 50–57. [CrossRef] [PubMed]
19. Bjørnskau, T. *Road Traffic Exposure and Risk among High-Risk Groups in Norway*; Insitute of Transport Economics Norwegian Centre for Transport Reserch: Oslo, Norway, 2009.
20. Elvik, R. Why some road safety problems are more difficult to solve than others. *Accid. Anal. Prev.* **2010**, *42*, 1089–1096. [CrossRef]
21. Liu, C.C.; Hosking, S.G.; Lenné, M.G. Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment. *Transp. Res. Part F Traffic Psychol. Behav.* **2009**, *12*, 325–334. [CrossRef]
22. Allen, T.; Newstead, S.; Lenné, M.G.; McClure, R.; Hillard, P.; Symmons, M.; Day, L. Contributing factors to motorcycle injury crashes in Victoria, Australia. *Transp. Res. Part F Traffic Psychol. Behav.* **2017**, *45*, 157–168. [CrossRef]
23. Jimenez, A.; Bocarejo, J.P.; Zarama, R.; Yerpez, J. A case study analysis to examine motorcycle crashes in Bogota, Colombia. *J. Saf. Res.* **2015**, *52*, 29–38. [CrossRef] [PubMed]
24. Simone, A.; Vignali, V.; Marinelli, M. Intersections design and powered two wheelers interaction: Experimental investigation of an accident prediction model. In Proceedings of the 4th International SIIV Congress, Palermo, Italy, 12–14 September 2007.
25. Yasin, Y.J.; Grivna, M.; Abu-Zidan, F.M. Motorized 2–3 wheelers death rates over a decade: A global study. *World J. Emerg. Surg.* **2022**, *17*, 7. [CrossRef]
26. Ijaz, M.; Liu, L.; Almarhabi, Y.; Jamal, A.; Usman, S.M.; Zahid, M. Temporal Instability of Factors Affecting Injury Severity in Helmet-Wearing and Non-Helmet-Wearing Motorcycle Crashes: A Random Parameter Approach with Heterogeneity in Means and Variances. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10526. [CrossRef]
27. Ijaz, M.; Lan, L.; Usman, S.M.; Zahid, M.; Jamal, A. Investigation of factors influencing motorcyclist injury severity using random parameters logit model with heterogeneity in means and variances. *Int. J. Crashworthiness* **2021**, *27*, 1412–1422. [CrossRef]
28. Santos, K.; Dias, J.P.; Amado, C.; Sousa, J.; Francisco, P. Risk factors associated with the increase of injury severity of powered two wheelers road accidents victims in Portugal. *Traffic Inj. Prev.* **2021**, *22*, 646–650. [CrossRef]
29. Ijaz, M.; Lan, L.; Zahid, M.; Jamal, A. A comparative study of machine learning classifiers for injury severity prediction of crashes involving three-wheeled motorized rickshaw. *Accid. Anal. Prev.* **2021**, *154*, 106094. [CrossRef] [PubMed]
30. Abdelwahab, H.T.; Abdel-Aty, M.A. Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. *Transp. Res. Rec. J. Transp. Res. Board* **2001**, *1746*, 6–13. [CrossRef]
31. Penmetsa, P.; Pulugurtha, S.S. Risk drivers pose to themselves and other drivers by violating traffic rules. *Traffic Inj. Prev.* **2017**, *18*, 63–69. [CrossRef] [PubMed]
32. Anvari, M.B.; Tavakoli Kashani, A.; Rabieyan, R. Identifying the Most Important Factors in the At-Fault Probability of Motorcyclists by Data Mining, Based on Classification Tree Models. *Int. J. Civ. Eng.* **2017**, *15*, 653–662. [CrossRef]

33. Duddu, V.R.; Penmetsa, P.; Pulugurtha, S.S. Modeling and comparing injury severity of at-fault and not at-fault drivers in crashes. *Accid. Anal. Prev.* **2018**, *120*, 55–63. [CrossRef]
34. Rezapour, M.; Mehrara Molan, A.; Ksaibati, K. Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models. *Int. J. Transp. Sci. Technol.* **2020**, *9*, 89–99. [CrossRef]
35. Lin, L.; Shi, F.; Li, W. Assessing inequality, irregularity, and severity regarding road traffic safety during COVID-19. *Sci. Rep.* **2021**, *11*, 13147. [CrossRef]
36. Bao, J.; Yang, Z.; Zeng, W.; Shi, X. Exploring the spatial impacts of human activities on urban traffic crashes using multi-source big data. *J. Transp. Geogr.* **2021**, *94*, 103118. [CrossRef]
37. Song, L.; Fan, W.; Li, Y. Time-of-day variations and the temporal instability of multi-vehicle crash injury severities under the influence of alcohol or drugs after the Great Recession. *Anal. Methods Accid. Res.* **2021**, *32*, 100183. [CrossRef]
38. Jiang, F.; Yuen, K.K.R.; Lee, E.W.M. Analysis of motorcycle accidents using association rule mining-based framework with parameter optimization and GIS technology. *J. Saf. Res.* **2020**, *75*, 292–309. [CrossRef]
39. Lee, Y.; Kim, D.-S. Internal Migration in South Korea. In *Internal Migration in the Countries of Asia*; Springer International Publishing: Cham, Switzerland, 2020; pp. 93–111.
40. Lee, S.; Noland, R.B. Bikeshare trips in Seoul, South Korea. *Case Stud. Transp. Policy* **2021**, *9*, 743–755. [CrossRef]
41. OECD. OECD Stat. Available online: https://stats.oecd.org/index.aspx?queryid=73639 (accessed on 21 June 2021).
42. Chung, Y.; Song, T.-J.; Yoon, B.-J. Injury severity in delivery-motorcycle to vehicle crashes in the Seoul metropolitan area. *Accid. Anal. Prev.* **2014**, *62*, 79–86. [CrossRef] [PubMed]
43. WHO (World Health Organization). Death on the Roads. Available online: https://extranet.who.int/roadsafety/death-on-the-roads/#country_or_area/GHA (accessed on 13 May 2022).
44. Kim, S.; Chung, Y. Injury Severity Analysis of Cyclists in Two Wheeler to Taxi Crashes: An Application of Vehicle Black Box Data in Incheon, Korea. *J. Korean Soc. Civ. Eng.* **2018**, *38*, 917–923.
45. Park, H.-C.; Joo, Y.-J.; Kho, S.-Y.; Kim, D.-K.; Park, B.-J. Injury Severity of Bus–Pedestrian Crashes in South Korea Considering the Effects of Regional and Company Factors. *Sustainability* **2019**, *11*, 3169. [CrossRef]
46. Mann, H.B. Nonparametric tests against trend. *Econometrica* **1945**, *13*, 245–259. [CrossRef]
47. Kendall, M.G.; Gibbons, J.D. *Rank Correlation Methods*, 5th ed.; Oxford University Press: London, UK, 1990.
48. Pande, A.; Abdel-Aty, M. Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. *Saf. Sci.* **2009**, *47*, 145–154. [CrossRef]
49. Hong, J.; Tamakloe, R.; Park, D. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. *Accid. Anal. Prev.* **2020**, *142*, 105497. [CrossRef]
50. Samerei, S.A.; Aghabayk, K.; Mohammadi, A.; Shiwakoti, N. Data mining approach to model bus crash severity in Australia. *J. Saf. Res.* **2021**, *76*, 73–82. [CrossRef]
51. Weng, J.; Zhu, J.-Z.; Yan, X.; Liu, Z. Investigation of work zone crash casualty patterns using association rules. *Accid. Anal. Prev.* **2016**, *92*, 43–52. [CrossRef]
52. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019. Available online: https://www.R-project.org/ (accessed on 4 May 2020).
53. Kwak, Y.-A.; Cho, Y.-S. Strategy of Food Retailer and Delivery Rider's Accident in South Korea. *J. Distrib. Sci.* **2020**, *18*, 49–60.
54. Tamakloe, R.; Park, D. Factors influencing fatal vehicle-involved crash consequence metrics at spatio-temporal hotspots in South Korea: Application of GIS and machine learning techniques. *Int. J. Urban Sci.* **2022**, 1–35. [CrossRef]
55. Hong, J.; Park, J.; Lee, G.; Park, D. Endogenous commercial driver's traffic violations and freight truck-involved crashes on mainlines of expressway. *Accid. Anal. Prev.* **2019**, *131*, 327–335. [CrossRef] [PubMed]
56. Kodaka, K.; Otabe, M.; Urai, Y.; Koike, H. Rear-End Collision Velocity Reduction System. Available online: http://www.jstor.org/stable/44745417 (accessed on 28 June 2021).
57. Haque, M.M.; Chin, H.C. Right-Angle Crash Vulnerability of Motorcycles at Signalized Intersections. *Transp. Res. Rec. J. Transp. Res. Board* **2010**, *2194*, 82–90. [CrossRef]
58. Byun, J.H.; Jeong, B.Y.; Park, M.H. Characteristics of Motorcycle Crashes of Food Delivery Workers. *J. Ergon. Soc. Korea* **2017**, *36*, 157–168.
59. Perez-Fuster, P.; Rodrigo, M.F.; Ballestar, M.L.; Sanmartin, J. Modeling offenses among motorcyclists involved in crashes in Spain. *Accid. Anal. Prev.* **2013**, *56*, 95–102. [CrossRef]
60. Park, J.S.; Jung, Y.I.; Kim, Y.H. Characteristics and Influencing Factors of Red Light Running (RLR) Crashes. *J. Korean Soc. Transp.* **2014**, *32*, 198–206. [CrossRef]
61. Campbell, B.N.; Smith, J.D.; Najm, W. *Analysis of Fatal Crashes Due to Signal and Stop Sign Violations*; U.S. Department of Transportation: Washington, DC, USA, 2004.