Accident Analysis and Prevention 165 (2022) 106517

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Factors affecting motorcycle crash casualty severity at signalized and non-signalized intersections in Ghana: Insights from a data mining and binary logit regression approach

Reuben Tamakloe [a], Subasish Das [b], Eric Nimako Aidoo [c], Dongjoo Park [d,*]

[a] Department of Transportation Engineering, The University of Seoul, 163 Seoulsiripdae-ro Dongdaemun-gu, Seoul 02504, South Korea
[b] Texas A&M Transportation Institute, College Station, TX 77843, USA
[c] Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[d] Department of Transportation Engineering & Department of Urban Big Data Convergence, The University of Seoul, 163 Seoulsiripdae-ro Dongdaemun-gu, Seoul 02504, South Korea

ARTICLE INFO

Keywords:
Developing countries
Motorcycle
Data mining
Safety
Intersection
Injury severity

ABSTRACT

Despite the countless benefits derived from motorcycle usage, it has become a significant public health concern, particularly in developing countries, due to the plateauing number of fatal/serious injuries associated with them. Although it has been well documented that the frequency and fatality rates of intersection-related motorcycle crashes are high, little research efforts have been made to explore the contributory factors influencing motorcycle-involved crashes at these locations. Interestingly, no study has investigated the latent patterns and chains of factors that simultaneously contribute to the injury severity sustained by motorcycle crash casualties at intersections under different traffic control conditions in developing countries. Since motorcycles are mostly used as taxis in developing countries, it is imperative to consider the injury severity sustained by all crash casualties in the motorcycle safety analysis. This study bridges the research gap by employing a plausible data mining tool to explore hidden rules associated with motorcycle crash casualty injury severity outcomes at both signalized and non-signalized intersections in Ghana's most densely populated region, Accra, using three-year crash data spanning 2016–2018. Besides, a binary logit regression model was also employed to explore the impact of crash factors on casualty severity outcomes using the same dataset. The results from both analysis techniques were consistent; however, the data mining technique provided chains of factors which provided additional insights into the groups of factors that collectively influence the casualty injury severity outcomes. From the rule discovery results, while full license status, daytime/daylight, and shoulder presence increased the risk of fatal injuries at signalized intersections, factors such as inattentiveness, good road surface, nighttime, shoulder absence, and young rider were highly likely to increase casualty fatalities at non-signalized intersections. By controlling all or some of these risk factors, the level of injury severity on the roadways could be reduced. Based on the findings, we provide enforcement, education, and engineering-based recommendations to help improve motorcycle safety.

1. Introduction

1.1. Background

Motorcycle usage is increasing globally, particularly in countries experiencing rapid economic growth. The main reason for this trend is that motorcycles provide riders with benefits in terms of easy maneuverability, low fuel consumption and operation cost, high travel speeds compared to cars, and less space requirement for parking (Jack et al., 2021). Nevertheless, the resort to using motorcycles as an alternative means of commute in these jurisdictions has introduced various dimensions of socioeconomic issues and public health concerns. The World Health Organization (WHO) noted that about 28% of the 1.35 million people lost to road traffic crashes yearly are motorcyclists (WHO, 2018). This trend is likely to increase, particularly in Africa, as the number of motorized vehicles continues to increase (WHO, 2017).

Although crashes can occur anywhere at any given time, research shows that motorcycle crashes occurring at intersections are likely to

* Corresponding author.
E-mail addresses: drtamakloe@uos.ac.kr (R. Tamakloe), s-das@tti.tamu.edu (S. Das), en.aidoo@yahoo.com (E. Nimako Aidoo), djpark@uos.ac.kr (D. Park).

https://doi.org/10.1016/j.aap.2021.106517
Received 13 August 2021; Received in revised form 27 November 2021; Accepted 29 November 2021
0001-4575/

have more fatal consequences than other vehicle users (Lin et al., 2004); (Nguyen et al., 2021); (Wahab and Jiang, 2019). Motorists are more exposed to traffic compared to the users of other vehicle types. Besides, the complex nature of urban intersections, coupled with the conflicting maneuvers and interactions between different traffic streams and the relatively low conspicuity at these locations, is identified to increase the vulnerability of motorcycle users (Haque et al., 2008); (Pai and Saleh, 2008); (Strauss et al., 2014). Despite the criticality and the high level of fatality associated with motorcycle crashes at these segments, it is not clear as to which chains of factors simultaneously influence the severity generation mechanism of these crashes.

In the literature, studies have explored the individual effects of risk factors on the fatality of injuries sustained by motorcycle crashes at intersections while focusing mainly on the rider's injuries (Nguyen et al., 2021); (Pai and Saleh, 2008; Pai and Saleh, 2007). By estimating separate models for different controls at intersections in the UK, it was identified that the injury severities sustained by motorcyclists at intersections differ significantly based on the type of traffic control present (Pai and Saleh, 2007). It is important to note that there have been diverging findings in studies that combined crashes irrespective of the control at the intersection. While some studies assert that the injuries sustained by motorcyclists at signalized intersections are less likely to be severe relative to those at non-signalized intersections due to the better safety environment present at the signalized intersection (Aidoo and Amoh-Gyimah, 2020), others showed that the presence of traffic signals increased the likelihood of fatalities, particularly in approach-turn cases (Haque et al., 2008); (Pai and Saleh, 2007); (Wahab and Jiang, 2019). Although these studies identified the impact of individual risk factors on motorcyclist (rider) injury severity outcomes considering the type of control present at the intersection, they did not provide information about which combinations of factors lead to each motorcycle crash casualty injury severity level. Besides, studies of this nature from developing countries that analyze separate models based on the control at intersections are nonexistent. Since motorcycles in most developing nations, particularly in Ghana, Togo, and Nigeria, are mainly used for transporting people (motor taxis popularly known as "okada") (Appiah, 2018), it is imperative to consider exploring the factors influencing the severity of injuries sustained not only by the motorcyclists (riders) but also all casualties involved in the crash.

To the best of our knowledge, research investigating the latent trends and combinations of factors influencing motorcycle safety at intersections, particularly regarding the type of control available at the intersection, is lacking despite their criticality and the high level of fatality associated with them. A better understanding of such multiple chains of factors and their patterns could be resourceful in providing a guide as to which combination of risk factors can be eliminated or suppressed to improve traffic safety. We posit that the results of this study will further facilitate the identification of targeted policy suggestions for improving motorcycle traffic safety in developing countries.

1.2. A brief review and critique of prior studies

This section briefly reviews prior research on factors influencing the severity of motorcycle crashes. It then focuses on studies exploring the influence of individual risk factors on motorcycle crashes at intersections considering the type of control present.

Research on crashes involving motorcycles is a topic that is gradually gaining traction. Over the years, this body of literature has dealt with issues influencing motorcycle safety in terms of crash risk and severity. Regarding rider features, a study conducted in Ghana has documented that the risk of motorcycle fatalities correlates with age (Aidoo and Amoh-Gyimah, 2020). According to a study from Greece, the effect of age on fatality risk decreases with age in the lower age groups and increases in the older age groups, placing 15 to 17-year-olds as the group with the highest risk (Yannis et al., 2005). Other studies have shown that gender, alcohol consumption, and the behavioral traits of riders do affect injury severity (Albalate and Fernández-Villadangos, 2010); (Vlahogianni et al., 2012). Concerning motorcycle type, engine size was identified to significantly influence severity, particularly in the higher capacity motorcycles (Yannis et al., 2005). Further, temporal and roadway-related features identified as having significant effects on motorcycle safety include weekend, median presence, darkness, and clear weather (Aidoo and Amoh-Gyimah, 2020); (Rifaat et al., 2012); (Wahab and Jiang, 2019).

Factors common to most developing countries have also been considered in some motorcycle safety studies. In Pakistan, female pillion riders were likely to be involved in fatal crashes, especially when their traditional loose and long clothes get stuck in motorcycles' chains or rear wheels (Pervez et al., 2021). (Wahab and Jiang, 2019) and (Wahab and Jiang, 2019) explored motorcycle crash data from Ghana containing factors relating to the road surface type and shoulder conditions; which according to them; were unique to the Ghanaian situation and have not been discussed in previous studies. The study identified that while overgrown shoulders and poor road surface condition (potholes) are more likely to cause injuries, crashes at intersections, at those at locations with traffic control presence (signage) are likely to result in fatalities. Other pertinent issues associated with fatal motorcycle crashes and are prevalent in most developing countries were also identified in Bangladesh, Tanzania, Ghana and India to be riding without a helmet and with an expired license (Nimako Aidoo et al., 2018); (Rahman et al., 2021); (Salum et al., 2019); (Sivasankaran et al., 2021).

Road infrastructure has been cited as having a strong link with fatal and severe injuries, particularly in developing countries. Besides, studies have alluded that providing adequate roadway infrastructure in terms of good roads and traffic signals and considering roadway design standards that consider all road users help make roads safer (Aidoo and Amoh-Gyimah, 2020); (Tamakloe et al., 2021); (WHO, 2018). Nevertheless, it is interesting to note that motorcycle crash studies focusing on the effects of control type at intersections (signalized/non-signalized) on severity outcomes in developing countries are limited.

In urban settings, many vulnerable road users are involved in crashes at intersections due to the complexities associated with them. Using a binary logit model, significant factors of motorcyclist injury severity at T-intersections were identified. According to UK-based researchers, non-signalized intersections were deadly as injury severities were greatest in crashes involving motorcycle going-ahead maneuvers and right-turning vehicles (Pai, 2009). Another study from the UK focused on investigating the risk factors influencing motorcyclist injury severity at three-legged intersections using an ordered probit model noted that for different types of control (Pai and Saleh, 2007). The study identified that whereas the collision partner, the crash month, light condition, weather, speed limit, engine size, and riders age had a similar effect on severity outcomes in terms of parameter sign, variables such as gender, day of the week, and time of day had diverging effects. It is noteworthy that, although these studies identified the effects of the risk factors on motorcyclist injury severity, they could not specify which groups of factors are likely to lead to each crash severity outcome. Besides, they did not consider the severity of all casualties involved in the motorcycle crash.

Recently, transportation safety experts have mostly resorted to using data mining techniques, particularly association rules mining (ARM), in exploring the chains of factors influencing crash severity. This technique is particularly advantageous as it explores interesting rules (rules with high correlations between antecedents and consequents) from datasets without predefined assumptions, as in the case of statistical methods (regression-based techniques) (Pande and Abdel-Aty, 2009). In the traffic safety literature, it has been extensively used in exploring trends and groups of factors leading to fatality in crashes involving pedestrians, general vehicles such as cars, buses, and trucks (Das et al., 2019; Das et al., 2021); (Hong et al., 2020; Hong et al., 2020); (Montella et al., 2011); (Pande and Abdel-Aty, 2009); (Samerei et al., 2021); (Weng et al., 2016). Despite its usefulness, few researchers have employed it in studying motorcycle crashes (Jiang et al., 2020); (Li and Wang, 2020); (Montella et al., 2020); Montella et al., 2012).

Concerning motorcycle crashes, a study conducted in Italy identified that crashes occurring in rural areas during the nighttime are likely to be fatal. Besides, a combination of curve alignments and run-off-road crashes were highly associated with fatalities. According to the study, other factors strongly correlated with motorcycle crash fatalities include truck involvement, head-on crash type, and rural road types (Montella et al., 2012). Spain-based researchers who investigated patterns among factors influencing powered two-wheeler (PTW) crashes mentioned that those very young riders involved in single-vehicle crashes during the daytime are likely to have fatal consequences. Besides, male motorcyclists involved in single-vehicle collisions when riding during the daytime while it rains are likely to have fatal injuries. The authors recommended using data mining tools to explore crash data to identify chains of factors as it provides less biased results since they do not rely on statistical assumptions (Montella et al., 2020). (Jiang et al., 2020) employed the ARM approach to explore critical risk factors in a motorcycle crash database containing 24; 680 crashes that occurred in Victoria, Australia, from 2006 to 2017. The study identified that motorcycle collision with trucks and fixed objects, crashes occurring late in the night or during early morning hours, overtaking, and collisions with vehicles moving in opposite directions are key factors likely to cause fatal injury. Besides, the authors identified that the key factors likely to influence non-fatal injury include motorcycle turning movements, being female, and non-collision crashes (no object struck and overturns). (Li and Wang, 2020) also explored the factors influencing mortality in single-motorcycle crashes in China using the ARM technique. From the results obtained; factors such as non-helmet usage, low-grade highways, straight section roads, sunny weather, and riders in their 45-50′s were associated with deaths.

1.3. Study objective and contribution

The reviewed literature shows that motorcycle safety researchers did not consider the discovery of rules influencing the injury severity sustained by the casualties involved in motorcycle crashes at both signalized and non-signalized intersections. As motorcycle usage and fatality rates in most nations, including developing countries, is on the rise (Wahab and Jiang, 2019; Wahab and Jiang, 2019); (WHO, 2017), identifying groups of influential risk factors that show strong associations with each crash casualty severity outcome under various traffic control measures at intersections in these nations is imperative and should be considered.

This study aims to bridge the research gap by (i) investigating the contributory factors in leading to motorcycle casualty severity outcomes at intersections and (ii) exploring and understanding the differences between the chains of factors in motorcycle crashes associated with each casualty severity outcome at both signalized and non-signalized intersections in a developing country. This objective is achieved by using the association rule mining (ARM) technique, which involves a clearly defined task of identifying clusters of crash-risk factors that occur together in a crash based on user-defined thresholds without relying on the numerous assumptions of parametric models, which may lead to biased estimates when violated (Teng, 2011). The severity of a crash can simply be reduced by reducing or eliminating at least one of the chains of factors obtained for each severity level (Samerei et al., 2021). As a benchmark, the binary logit regression model is also used to estimate the impact of risk factors on motorcycle crash casualty severity. Unlike the studies discussed above, which identify the individual impact of risk factors on motorcycle crash casualty injury severity, the effective chains of factors derived by ARM in this study can be used by transportation engineers in developing countries for the formulation of targeted policies geared at reducing the severity of motorcycle crashes.

2. Data description

2.1. The magnitude of motorcycle safety problem in Accra, Ghana

The study area for this analysis is the Greater Accra Region (GAR), the capital of Ghana. With a population of about five million and a land area of 3,245 km^2 or 1,253 mi^2, the GAR is the most densely populated region in the country. Due to the heavy traffic congestion and poor quality of service provided by public transport providers, many people have resorted to using motorcycles for inter and intra-city transport (Armah et al., 2010); (Jack et al., 2021). Further, the lack of viable public transport and the high unemployment rates has led to the rapid growth in privately owned unregulated motorcycle transport businesses in the GAR (Boateng, 2021); (Jack et al., 2021) (see Fig. 1. Despite the advantages of this mode of transport in terms of maneuverability, low operation cost, and comparable speeds to cars, it is mainly associated with the most fatal crashes on the roadways (Aidoo and Amoh-Gyimah, 2020). Researchers from Ghana highlight that motorcycle crashes are primarily due to the disregard of traffic safety regulations and the poor enforcement of existing regulations (Jack et al., 2021). According to a survey conducted by the authors, it was identified that motorcyclists fail to observe traffic regulations due to pressure by passengers, police harassment, the urge to meet deadlines, and sales intent, in that order. The majority of the respondents alluded to running red lights at the sight of a police officer due to police extortions and arrests. According to statistics, the number of motorcycle fatalities follows an upward trend as new motorcycle registrations keep on increasing. In particular, motorcycle-related deaths rose from 2.7% in 2001 to 35.3% in 2016 (MOT, 2020), and the chance of dying from a crash stands at a troubling ten times higher than that of a passenger vehicle crash (Boateng, 2021). This, together with other vehicular crashes, places GAR as the region with the highest number of crashes and the second region with the highest number of fatalities in Ghana (NRSC, 2016). The increasing trend in fatal/serious injuries is worrying, and the economic burden of crashes calls for a detailed analysis of the factors influencing motorcycle crash severity at critical crash locations.

2.2. Study data

Police-reported data extracted from the National Road Traffic Accident database at the Building and Road Research Institute of the Council for Scientific and Industrial Research were used for this study. The crashes relate to all motorcycle-involved collisions in the GAR from 2016 to 2018, and were collected at the casualty level. In total, 260 observations of intersection-related motorcycle crash casualties recorded in GAR within the period were used for the analysis. Data from the GAR was used for this study as motorcycles have gained grounds as one of the main modes of transport (Boateng, 2021); (Jack et al., 2021).

The data for this study, provided in Table 1 below, consists of comprehensive information relating to the crash characteristics (maneuver before crash, traffic violations, number of vehicles involved, and collision type), roadway features (road alignment, road surface condition, shoulder/median presence, intersection type, and traffic control at the intersection), environmental conditions (weather, and light condition), temporal features (day and time of the crash), vehicle characteristics (vehicle ownership), and casualty (age, sex, and injury severity sustained) and rider (license status) records. In this study, the casualty could either be the motorcycle rider or the pillion passenger. Crash casualty injury severity in Ghana is classified into three categories: K-fatal/killed (where death is recorded within 30 days of the crash), SI-hospitalized/severe/serious injury (where the casualty is hospitalized for at least 24 h), and MINJ-injured not-hospitalized/minor injury (where the crash victim sustains a little injury which may require less than 24 h of medical attention) (Amoh-Gyimah et al., 2017). Similar to previous studies, the observations relating to fatal/killed and hospitalized/severe injury were combined due to the low proportion of fatal

(a) A section of private commercialized motorcycle owners waiting for passengers (https://ghana.dubawa.org/okada-pix-1/)

(b) Motorcycle riders and passengers without helmets utilizing the road's shoulders (https://www.modernghana.com/news/1030179/okada-fight-bawumia-has-lost-touch-with-reality.html)

Fig. 1. Motorcycle problem in Ghana.

incidents in the dataset (9.6%). Thus, two injury severity levels, KSI and MINJ, were considered in this investigation.

3. Methodology

The focus of this study is to employ a descriptive-analytic approach known as the association rule mining (ARM) technique to extract valuable information from the motorcycle crash dataset. To also identify the individual impact of factors affecting motorcycle crash severity at signalized and non-signalized intersections as a benchmark, the traditional binary logistic regression approach is also employed. A brief explanation of the binary logit regression (BLR) model and the description of the main data mining technique (ARM) used for this study are provided in this section.

3.1. Binary logit regression model

The binary logit regression (BLR) has been used in traffic safety for analyzing the relationship between outcome variables of a binary nature and explanatory variables (Young and Liesman, 2007). In this study, the outcome variable y_{in} is of a dichotomous nature defined as:

$$y_{in} = \begin{cases} 1, & \text{if the severity outcome } i \text{ sustained by observation } n \text{ was } KSI \\ 0, & \text{if the severity outcome } i \text{ sustained by observation } n \text{ was } MINJ \end{cases} \quad (1)$$

If the probability that a motorcycle crash casualty n would be killed or severely injured (KSI) is $P(y_{in} = 1)$ then, the logistic function, as presented by (Manski and McFadden, 1981) can be defined as follows:

$$P(y_{in} = 1) = \frac{exp(\beta' X_{in})}{1 + exp(\beta' X_{in})} \quad (2)$$

where X_{in} is the vector of independent variables and β is the vector of estimable coefficients. Estimation of the models coefficients could be obtained by maximizing the log-likelihood function below (Nasri and

Table 1
Summary of variables used in the study.

Variables	Total data (260 obs.)		SIG (135 obs.)		NSIG (125 obs.)	
	N	%	N	%	N	%
Crash characteristics						
Number of vehicles involved						
Number of vehicles: single-vehicle crash	35	13	17	13	18	14
Number of vehicles: two-vehicle crash	206	79	105	78	101	81
Number of vehicles: three or more vehicle crash	19	7	13	10	6	5
Vehicle maneuver before crash						
Maneuver: cross traffic	8	3	3	2	5	4
Maneuver: going-ahead	226	87	117	87	109	87
Maneuver: left turn	11	4	8	6	3	2
Maneuver: merging	5	2	3	2	2	2
Maneuver: overtaking	5	2	3	2	2	2
Maneuver: right turn	2	1	1	1	1	1
Maneuver: U-turn	3	1	n.a.	n.a.	3	2
Collision type						
Collision type: head-on	22	8	9	7	13	10
Collision type: hit object/parked vehicle	9	3	3	2	6	5
Collision type: hit pedestrian	26	10	16	12	10	8
Collision type: ran off road/overturn	5	2	n.a.	n.a.	5	4
Collision type: rear-end	87	33	53	39	34	27
Collision type: right-angle	70	27	35	26	35	28
Collision type: sideswipe	41	16	19	14	22	18
Crash location characteristics						
Road shoulder						
Shoulder: none	122	47	80	59	42	34
Shoulder: present	138	53	55	41	83	66
Road separation						
Median: none	98	38	40	30	58	46
Median: present	162	62	95	70	67	54
Location (intersection) type						
Intersection type: T-intersection	148	57	61	45	87	70
Intersection type: Y-intersection	3	1	n.a.	n.a.	3	2
Intersection type: four-legged	96	37	73	54	23	18
Intersection type: roundabout	10	4	n.a.	n.a.	10	8
Intersection type: staggered intersection	3	1	1	1	2	2
Traffic control						
Control: non-signalized	125	48	n.a.	n.a.	n.a.	n.a.
Control: signalized	135	52	n.a.	n.a.	n.a.	n.a.
License status and violations						
License status						
License status: full	62	24	50	37	12	10
License status: provisional	1	0	n.a.	n.a.	1	1
License status: unknown	194	75	83	61	111	89
License status: unlicensed	3	1	2	1	1	1
Traffic violations/error						
Violation: improper turning	3	1	2	1	1	1
Violation: inattentive	92	35	50	37	42	34
Violation: inexperience	6	2	5	4	1	1
Violation: none	92	35	43	32	49	39
Violation: other (improper turning/fatigue/improper overtaking/no signal)	13	5	6	4	7	6
Violation: too close	10	4	1	1	9	7
Violation: too fast	37	14	22	16	15	12
Violation: unknown	7	3	6	4	1	1

Table 1 (*continued*)

Variables	Total data (260 obs.)		SIG (135 obs.)		NSIG (125 obs.)	
Temporal characteristics						
Day of week						
Day: weekday	183	70	93	69	90	72
Day: weekend	77	30	42	31	35	28
Time of day						
Time: daytime	176	68	90	67	86	69
Time: nighttime	84	32	45	33	39	31
Environmental characteristics						
Weather condition						
Weather: clear	240	92	119	88	121	97
Weather: inclement	20	8	16	12	4	3
Light conditions						
Light condition: daylight	184	71	91	67	93	74
Light condition: night-lights off	2	1	n.a.	n.a.	2	2
Light condition: night-lights on	72	28	43	32	29	23
Light condition: night-no lights	2	1	1	1	1	1
Roadway characteristics						
Road description						
Alignment: curve and inclined	1	0	1	1	n.a.	n.a.
Alignment: straight and flat	259	100	134	99	125	100
Road surface type						
Surface condition: good	257	99	135	100	122	98
Surface condition: potholes/rough	3	1	n.a.	n.a.	3	2
Vehicle characteristics						
Vehicle ownership/usage						
Ownership: other (government/emergency/police/military/company)	26	10	9	7	17	14
Ownership: private	234	90	126	93	108	86
Casualty characteristics						
Sex						
Sex: female	22	8	11	8	11	9
Sex: male	237	91	123	91	114	91
Sex: unknown	1	0	1	1	n.a.	n.a.
Age						
Age: from 30 to 50	116	45	59	44	57	46
Age: >50	14	5	6	4	8	6
Age: less than 30	113	43	59	44	54	43
Age: unknown	17	7	11	8	6	5
Severity level						
Severity: KSI	140	54	72	53	68	54
Severity: MINJ	120	46	63	47	57	46

Note: SIG: motorcycle-involved crash at a signalized intersection; NSIG: motorcycle-involved crash at a non-signalized intersection; N: number of crash casualty observations

Aghabayk, 2021):

$$LL(\beta) = \sum_{i=1}^{n}\{ (y_{in} ln(P(y_{in})) + (1 - y_{in}) ln(1 - P(y_{in})))\} \quad (3)$$

Marginal effect estimates, which shows how the probabilities change as a dummy regressor changes from 0 to 1 or as a continuous variable change by a single unit change, are computed to better explain the results of the fitted model. The logistic procedure used for obtaining the parameter coefficients and the marginal effect estimates were carried out using STATA 16. All variables were initially tested for significance, and, following previous studies (Tamakloe et al., 2021), those variables that do not have a significant relationship with the outcome variable were excluded from the model.

3.2. Association rule mining

ARM was first introduced by (Agrawal et al., 1993) and was initially used in discovering association rules in a large database of customer transactions. Since then; it has been broadly accepted and used by transportation researchers for discovering latent rules highlighting crash-contributory factors that frequently occur together in a dataset (Das et al., 2019); (Hong et al., 2020). The main advantage of the ARM is that it focuses on identifying hidden patterns in data rather than confirming a hypothesis, and it does not require prespecified assumptions. Besides, it performs well even in the case of missing data (Das et al., 2019), thus reducing the tendency to draw biased and inconsistent conclusions, which could suggest erroneous countermeasures. Further, it can handle both small and large datasets, making it suitable for analyzing rare events data. It is noteworthy that in previous transport safety research, ARM has been used to investigate chains of crash-contributory factors in small datasets of sample size ranging from 126 to 449 observations which otherwise could not be possible using regression-based statistical techniques due to sample size-related assumptions (Das et al., 2020); (Hong et al., 2020); (Montella, 2011); (Xu et al., 2018).

Suppose the intersection-related motorcycle crash casualty database is $M = \{m_1, m_2, \cdots, m_y\}$, such that each casualty observation is made up of a subset of y unique crash-risk factors or items present in the itemset Q , such that, $Q = \{q_1, q_2, \cdots, q_y\}$. Then, association rules mined by the ARM methodology can be formulated as an implication of the form $\{A => B\}$, such that $A, B \subseteq Q$, and $A \cap B = \varnothing$. A , which is the left-hand side (LHS) of the rule, is known as the antecedent, and B on the right-hand side is the consequent of the rule. Thus, the rule $\{A => B\}$ can be described as: if A occurs or exists as a crash-risk factor, then B is expected to occur.

The framework proposed by (Agrawal et al., 1993) employed two parameters; namely, support and confidence, for evaluating the rules generated. The support, S , of a rule depicts how prevalent the itemset is. It does this by expressing the proportion of the entire database covered by the rule. Supports are be expressed as follows:

$$S(A => B) = \frac{|A \cap B|}{|M|}; S(A) = \frac{|A|}{|M|}; S(B) = \frac{|B|}{|M|} \quad (4)$$

where $S(A => B)$ refers to the support of the rule $A => B$, $S(A)$ is the support of the antecedent, $S(B)$ connotes the consequents support, $|A \cap B|$ is the frequency at which itemsets A and B co-occur, $|A|$ represents the number of times A occurs in the database, $|B|$ represents the number of times B occurs in the database, and $|M|$ represents the number of motorcycle crashes in the database (Montella et al., 2021). A rule's confidence, C , shows the percentage of cases in which the consequent, B , occurs given that the antecedent, A , has occurred. Higher values of C depicts that the presence of the rule's consequent B is high among the observations with the antecedent A. Confidence of the rule A → B is estimated as:

$$C(A => B) = \frac{S(A => B)}{S(A)} \quad (5)$$

The support–confidence framework was complemented by the measure "interest" or "lift", which efficiently prunes infrequent itemsets from the dataset (Brin and Motwani, 1998). The lift L of a rule A → Brelates how frequently the antecedent and consequent co-occur. It can also be described as how likely an itemset B occurs given that A has occurred, while controlling for the popularity of itemset B (Montella et al., 2020); (Samerei et al., 2021). It is computed as follows:

$$L(A => B) = \frac{S(A => B)}{S(A) \cdot S(B)} \quad (6)$$

A rule is deemed "interesting", "valuable", and is more preferred when its lift is>1, indicating strong interdependence between both items A and B , connoting that the associations did not occur by chance (Pande and Abdel-Aty, 2009); (Samerei et al., 2021). In other words, it shows how often the two items are a part of the same observation (Pande and Abdel-Aty, 2009). It is noteworthy that higher lift values connote higher "interestingness" of the rule (Hong et al., 2020). Besides, there is no correlation between both A and B when the lift is 1. A negative interdependence is established when the lift of the rule is less than 1.

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. Following previous studies, appropriate threshold values are primarily obtained after conducting a trial-and-error experiment (Das et al., 2019); (Hong et al., 2020). However, to ensure that redundant rules are further removed from the set of rules mined, transport safety researchers of recent recommended the validity verification of the addition of each item q to the rule. This ensures that the addition of a new item results in a sufficient increase in the lift (Montella et al., 2020; Montella et al., 2021); (Samerei et al., 2021). To do this, a rule with one item q in the antecedent itemset having a lift value L_{A_q} is used as the parent rule. The lift value, $L_{A_{q+1}}$, obtained after adding a new item $q+1$ is then compared to the previous lift value Lift_{A_q} using the LIC formula as follows:

$$LIC = \frac{L_{A_q}}{Lift_{A_{q+1}}} \quad (7)$$

where A_q is the antecedent with item q , and A_{q+1} is the antecedent after a new item is added. A rule with LIC greater than the preselected minimum threshold is more favored (Samerei et al., 2021).

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 were selected and employed for analysis: $s = 1\%$, $c = 5\%$, $l = 1.10$, and $LIC = 1.02$ (Montella et al., 2020); (Samerei et al., 2021).

4. Results and discussions

4.1. Preliminary findings

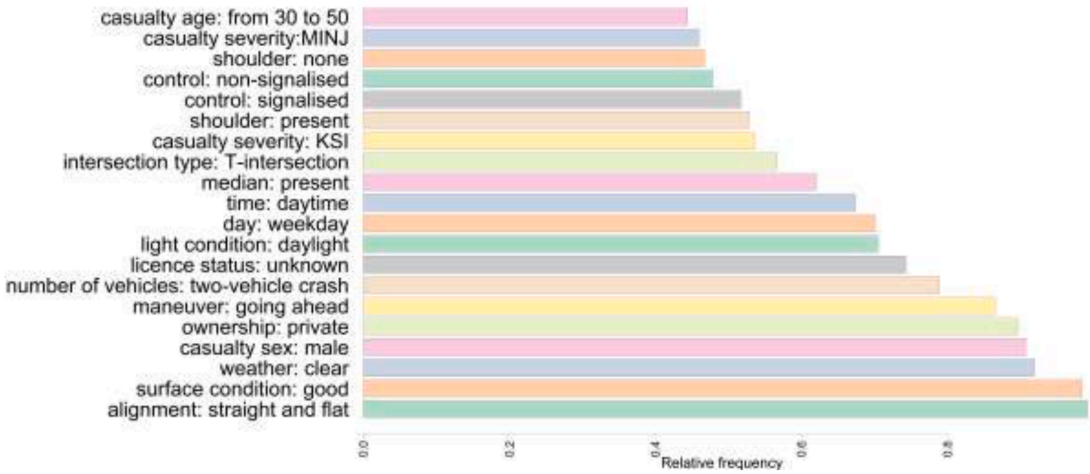
The generation of association rules was conducted using the R open-source data mining tool (R Core Team, 2017). Prior to the rule mining, the top 20 frequent items influencing motorcycle crashes were identified and presented in Fig. 2. From the figure, it is observed that the four most frequent items common to crashes at both signalized and non-signalized intersections relate to roadway geometry (straight and flat) and condition (good), vehicle ownership (private), and weather (clear). The item for motorcycle crashes at four-legged intersections was only present in the top items for crashes in signalized intersections (Fig. 2 (b)). Further, crashes at shoulder-absent non-signalized and median-absent signalized intersections were infrequent (not part of the items in Fig. 2b and Fig. 2c.

4.2. Findings from the binary logit regression model

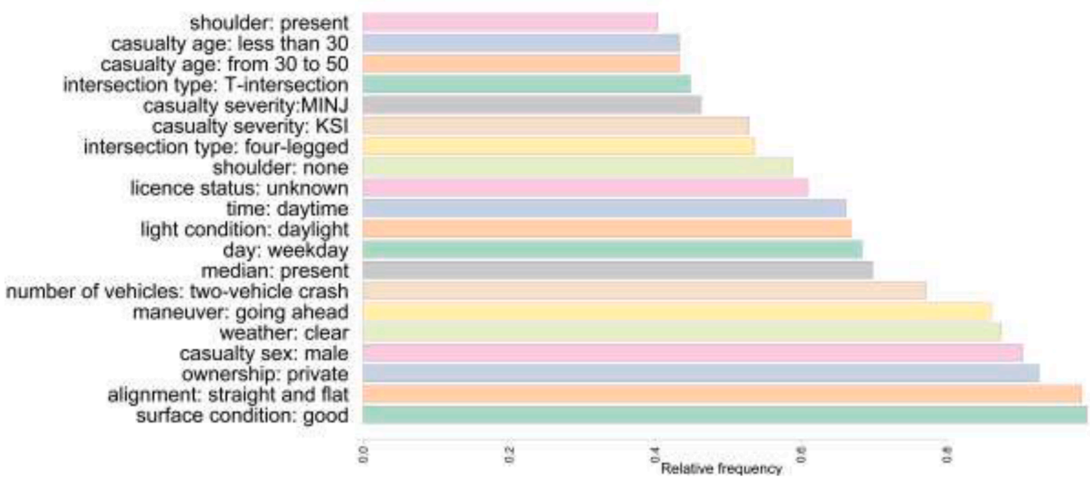
4.2.1. Crash characteristics variables

Variables categorized under crash characteristics were significant only in the model for the signalized and non-signalized intersection crash datasets. Regarding the signalized intersection crash model, it was identified that sideswipes and collisions involving three or more vehicles reduced the likelihood for KSI. In particular, the marginal effect estimates in Table 2 show that these variables reduced the probability for KSI by −0.427 and −0.337, respectively, if the crash occurs at a signalized intersection. These findings are intuitive as drivers/riders are likely to make turning or lane changing maneuvers more thoughtfully at these intersections due to the presence of traffic control. Besides, as riders in Ghana often weave through traffic during the red traffic signal phase to wait at the stop line (Aidoo and Amoh-Gyimah, 2020), the sideswipes that occur at these segments will likely cause minor injuries or damages. On the other hand, multi-vehicle collisions are likely to impact motorcyclists more, especially when drivers/riders exceed the posted speed limits. However, the finding from this study is likely as most roads in GAR are mostly congested, resulting in reduced speed limits (Agyapong and Ojo, 2018); (Møller-Jensen, 2021). Thus, multi-vehicle crashes at intersections may be caused by inattentiveness and the high interaction between different traffic streams, making it less likely to observe KSI severity outcomes.

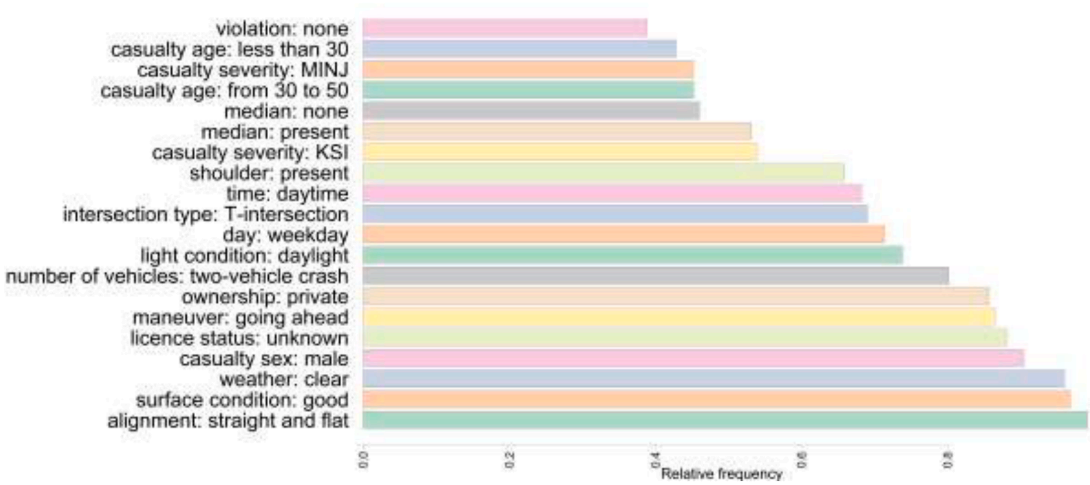
In the case of the non-signalized intersection, the indicator for the right-angled crash is found to increase the likelihood of KSI severity



(a) Frequent items associated with motorcycle crashes (full data)



(b) Frequent items associated with signalized intersection-related motorcycle crashes



(c) Frequent items associated with non-signalized intersection-related motorcycle crashes

Fig. 2. Frequent items in motorcycle crash data.

Table 2
Binary logit model regression results.

Variables	Total				Signalized intersection				Non-signalized intersection			
	Coef.	Std. Error	P-value	Marginal effect	Coef.	Std. Error	P-value	Marginal effect	Coef.	Std. Error	P-value	Marginal effect
Constant	−0.402	0.305	0.188	–	0.779	0.280	0.005	–	−0.347	0.765	0.651	–
Crash characteristics												
Collision type (reference: head-on)												
Collision type: right-angle	–	–	–	–	–	–	–	–	0.942	0.507	0.063	0.221
Collision type: sideswipe	–	–	–	–	−1.988	0.832	0.017	−0.427	–	–	–	–
Number of vehicles involved (reference: two-vehicle crash)												
Num. of vehicles: three or more vehicle crash	–	–	–	–	−1.462	0.747	0.050	−0.337	–	–	–	–
Crash location characteristics	–	–	–	–	–	–	–	–	–	–	–	–
Road separation (reference: median present)												
Median: none	–	–	–	–	–	–	–	–	−0.965	0.425	0.023	−0.234
Location (intersection) type (reference: staggered intersection)												
Intersection type: four-legged	–	–	–	–	–	–	–	–	1.119	0.609	0.066	0.252
Intersection type: T-intersection	−0.433	0.262	0.099	−0.107	–	–	–	–	–	–	–	–
License status and violations												
License status (reference: full)												
License status: unknown	−0.563	0.305	0.065	−0.137	–	–	–	–	−1.551	0.716	0.030	−0.318
Traffic violations/error (reference: too fast)												
Violation: inattentive	–	–	–	–	−0.906	0.465	0.052	−0.223	–	–	–	–
Temporal characteristics												
Day of week (reference – day: weekend)												
Day: weekday	–	–	–	–	–	–	–	–	1.095	0.498	0.028	0.267
Casualty characteristics												
Age (reference: from 30 to 50)												
Casualty age: less than 30	0.766	0.266	0.004	0.187	–	–	–	–	1.058	0.423	0.012	0.253
Sex (reference – sex: female)												
Sex: male					1.293	0.774	0.095	0.303				
Vehicle characteristics												
Vehicle ownership/usage (reference: private)												
Ownership: other	1.050	0.475	0.027	0.237	–	–	–	–	–	–	–	–
Model statistics												
Number of observations	260				135				125			
Log likelihood	−170.356				−72.562				−61.038			
Pseudo R2	0.244				0.231				0.158			

outcomes. In particular, when a right-angled crash occurs at a non-signalized intersection, the propensity for KSI increases by 0.221. Intuitively, right-angle crashes at these intersections in GAR are likely to be KSI as most motorcycle riders, predominantly motorcycle taxi operators or delivery workers, are fond of disregarding the right-of-way rules to reach their destinations in time (Boateng, 2021). This careless behavior exhibited at non-signalized intersections exposes them to unsuspecting road users, causing crashes with devastating severity outcomes.

4.2.2. Crash location characteristics variables

In terms of the crash location, the indicator variable for crashes occurring at T-intersections was significant only in the full data model. In this model, a crash at a T-intersection is less likely to be KSI as the probability for observing KSI outcomes decreases by −0.107. The study also identified that the probability of being killed or sustaining severe injuries increases by 0.252 when the crash occurs at a non-signalized four-legged intersection but decreases by −0.234 when it occurs at a non-signalized intersection without median separation. Four-legged intersections are complex and expose motorcyclists to different traffic streams relative to other intersection types (crossroads and roundabouts). Previous studies noted that crashes are more common at these intersections than others (Pai and Saleh, 2007). The lack of traffic signals further increases the complexity of these intersections. For these reasons, KSI outcomes are likely to be observed in crashes when road users fail to observe traffic rules and drive inattentively, which is a significant safety issue in Ghana (Boateng, 2021). The finding regarding median absence seems counterintuitive, as medians reduce traffic conflicts, resulting in fewer and less severe crashes. Nevertheless, this finding may be likely as, by design, median separated roads are wider, giving road users more room to operate (Kim et al., 2007). The elevated feeling of safety associated with good/wide roads, coupled with the indiscipline behavior of most riders in the GAR, may cause these road users to become less careful as they use non-signalized intersections, causing severe injuries.

4.2.3. License status and violation variables

Relative to crashes where the vehicle operator had a full license, those whose license status was unknown were found to have a reduced probability for KSI in both the full model and the model for crashes at non-signalized intersections. The probability for KSI in the full model was reduced by −0.137, and that of the non-signalized intersection model was reduced by −0.318. Most motorcycle riders acquire licenses through dubious means (aided mainly by middlemen or "Goro boys" at the license-providing agencies) without attending driving school. By doing so, they circumvent essential training and requirements such as eye tests, in-traffic tests, and physical examinations prescribed by Ghana's Driver and Vehicle Licensing Authority (DVLA) (Obiri-Yeboah et al., 2021). Thus, many motorcyclists have illegal full license status and frequently indulge in risky behaviors resulting in severe safety ramifications. Previous studies also showed that these riders are mostly involved in fatal crashes (Salum et al., 2019).

Regarding traffic violations, only the variable for inattentive driving was significant at 90% confidence level. In particular, our results showed that, compared to driving too fast/overspeeding, inattentive driving reduced the probability for KSI by −0.223. Although inattentive riding could cause crashes at signalized intersections, their severity could be reduced compared to overspeeding-related crashes. Previous studies also showed that speeding resulted in higher motorcycle crash severity outcomes (Pai and Saleh, 2007).

4.2.4. Temporal characteristics variables

Variables categorized under temporal characteristics were explored. Interestingly, the indicator variable for weekdays was the only significant variable in this category. The results showed that a crash at a non-signalized intersection during the weekday is likely to result in KSI relative to a crash that occurs on the weekend. From the marginal effects estimates, it is clear that the probability for KSI increases by 0.267 when a crash occurs during the weekday at a non-signalized intersection. This finding is plausible as riders are likely to drive fast to work or run errands during the weekdays when business is brisk, particularly at non-signalized intersections due to the reduced delays from the absence of traffic lights. This makes most motorcyclists pay less attention to the right-of-way rules. This act increases the level of confusion and the severity level in case of a crash.

4.2.5. Casualty characteristics variables

Under the casualty characteristics category, significant variables were identified for casualty age and sex indicators. Regarding the casualty age, it was observed that crash casualties of age less than 30 were likely to be killed or sustain severe injuries in the event of a motorcycle crash relative to those aged from 30 years to 50 years. This variable was significant in the full model and the non-signalized intersection crash model. Based on the marginal effects estimates, the probability for KSI is found to increase by 0.187 in the full model and by 0.253 in the non-signalized intersection model. Concerning the casualty sex, males are identified to be likely involved in KSI compared to females. The propensity for KSI crashes increases by 0.303 if the crash casualty is male. This variable was only significant in the signalized intersection model. Both young motorcyclists and males are likely to indulge in risky riding and are less likely to obey traffic regulations. Studies from Ghana show that, relative to males, females are more likely to own a valid motorcycle license and ride with a helmet (Nimako Aidoo et al., 2018). For these reasons peculiar to males and young riders, they will likely sustain KSI severity outcomes in crashes.

4.2.6. Vehicle characteristics variables

In terms of vehicle characteristics, the study identified that, relative to crashes involving vehicles used for private purposes, those used for “other” purposes (government/emergency/police/military/company) are likely to increase the likelihood of KSI if they crash with a motorcycle. From the marginal effects estimates, the propensity for KSI outcome increases by 0.237 when a crash occurs at an intersection (full model). This variable was insignificant in the other models (signalized and non-signalized intersections). This finding is likely as privately owned vehicle users are likely to drive carefully compared to the “other” group. Besides, as the “other” vehicle ownership group comprises mainly critical service providers, it is likely that they drive faster. Thus, motorcycles crashing with them at intersections are likely to be KSI.

4.3. Findings from the association rules mining technique

Association rules for motorcycle crashes in GAR, Ghana (irrespective of the intersection control type) are presented in Table 3. Besides, rules for motorcycle crashes at both signalized and non-signalized intersections are reported in Table 4 and Table 5, respectively. As displayed in the tables, the association discovery identified several strong rules for casualty severity category KSI and MINJ crashes based on the minimum thresholds (S > 10%, C > 50%, Lift > 1.05, and LIC > 1.10). The high lift levels show that the level of interdependency among the crash-contributory factors is strong. Notably, the lift values increased with every addition of a new item to the previously existing itemset. The number of rules generated for the various subsets of data studied under

Table 3
List of interesting rules for motorcycle crashes at intersections

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
Casualty severity: KSI	1–1	collision type: right-angle	18.01	67.14	1.25	47	n.a.
	1–2	collision type: right-angle, intersection type: four-legged	11.11	72.50	1.35	29	1.08
	1–3	casualty age: less than 30	27.20	62.83	1.17	71	n.a.
	1–4	casualty age: less than 30, intersection type: four-legged	11.88	70.45	1.31	31	1.12
	1–5	casualty age: less than 30, collision type: right-angle	10.34	69.23	1.29	27	1.10
	1–6	casualty age: less than 30, control: non-signalized	14.18	68.52	1.28	37	1.09
	1–7	casualty age: less than 30, control: non-signalized, number of vehicles: two-vehicle crash	13.03	73.91	1.38	34	1.08
	1–8	casualty age: less than 30, shoulder: present	16.09	67.74	1.26	42	1.08
	1–9	casualty age: less than 30, number of vehicles: two-vehicle crash	24.90	67.01	1.25	65	1.07
	1–10	casualty age: less than 30, median: present	17.24	66.18	1.23	45	1.05
	1–11	intersection type: four-legged	22.61	61.46	1.15	59	n.a.
	1–12	intersection type: four-legged, shoulder: present	12.26	65.31	1.22	32	1.06
	1–13	license status: full	14.56	61.29	1.14	38	n.a.
	1–14	license status: full, maneuver: going-ahead	13.41	64.81	1.21	35	1.06
Casualty severity: MINJ	2–1	casualty age: from 30 to 50	24.52	55.17	1.20	64	n.a.
	2–2	casualty age: from 30 to 50, license status: unknown	20.69	65.06	1.42	54	1.18
	2–3	casualty age: from 30 to 50, license status: unknown, shoulder: present	13.03	73.91	1.61	34	1.14
	2–4	casualty age: from 30 to 50, shoulder: present	14.18	62.71	1.36	37	1.14
	2–5	casualty age: from 30 to 50, control: non-signalized	13.03	59.65	1.30	34	1.08

(continued on next page)

Table 3 (*continued*)

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
	2–6	casualty age: from 30 to 50, control: non-signalized, license status: unknown	12.64	64.71	1.41	33	1.08
	2–7	casualty age: from 30 to 50, control: non-signalized, maneuver: going-ahead	13.03	62.96	1.37	34	1.06
	2–8	casualty age: from 30 to 50, time: nighttime	10.34	58.70	1.28	27	1.06
	2–9	collision type: rear-end	17.24	51.72	1.13	45	n.a.
	2–10	collision type: rear-end, shoulder: present	10.34	60.00	1.31	27	1.16
	2–11	collision type: rear-end, intersection type: T-intersection	10.73	56.00	1.22	28	1.08
	2–12	collision type: rear-end, time: daytime	12.64	55.93	1.22	33	1.08
	2–13	collision type: rear-end, light condition: daylight	12.64	55.93	1.22	33	1.08
	2–14	intersection type: T-intersection	28.74	50.68	1.10	75	n.a.
	2–15	intersection type: T-intersection, shoulder: present	17.62	54.76	1.19	46	1.08
	2–16	intersection type: T-intersection, violation: none	10.73	53.85	1.17	28	1.06
	2–17	day: weekend	14.94	50.65	1.10	39	n.a.
	2–18	day: weekend, license status: unknown	11.49	57.69	1.25	30	1.14
	2–19	day: weekend, license status: unknown, ownership: private	10.73	60.87	1.32	28	1.06
	2–20	day: weekend, shoulder: present	10.34	56.25	1.22	27	1.11
	2–21	day: weekend, ownership: private	14.18	55.22	1.20	37	1.09

the KSI category are as follows (full data (2-item rules: 26; 3-item rules: 261; 4-item rules: 1214), signalized intersection data (2-item rules: 23; 3-item rules: 220; 4-item rules: 933), non-signalized intersection data (2-item rules: 22; 3-item rules: 206; 4-item rules: 997)). Besides, the number of rules for the subset data explored under the MINJ category are as follows (full data (2-item rules: 6; 3-item rules: 85; 4-item rules: 419), signalized intersection data (2-item rules: 7; 3-item rules: 97; 4-item rules: 513), non-signalized intersection data (2-item rules: 7; 3-item rules: 88; 4-item rules: 474)).

Since we seek to obtain rules containing antecedents leading to a particular severity level, the association rules are mined for each subgroup by fixing the motorcycle crash casualty severity level as the

Table 4
List of interesting rules for motorcycle crashes at signalized intersections

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
Casualty severity: KSI	3–1	collision type: right-angle	16.18	62.86	1.19	22	n.a.
	3–2	collision type: right-angle, day: weekday	10.29	70.00	1.32	14	1.11
	3–3	license status: full	22.06	60.00	1.13	30	n.a.
	3–4	license status: full, violation: none	10.29	70.00	1.32	14	1.17
	3–5	license status: full, time: daytime	13.24	69.23	1.31	18	1.15
	3–6	license status: full, light condition: daylight	13.24	69.23	1.31	18	1.15
	3–7	license status: full, light condition: daylight, maneuver: going-ahead	12.50	77.27	1.46	17	1.12
	3–8	license status: full, median: present	15.44	65.63	1.24	21	1.09
	3–9	license status: full, median: present, number of vehicles: two-vehicle crash	12.50	73.91	1.40	17	1.13
	3–10	license status: full, shoulder: present	10.29	63.64	1.20	14	1.06
Casualty severity: MINJ	4–1	collision type: rear-end	20.59	52.83	1.14	28	n.a.
	4–2	collision type: rear-end, intersection type: T-intersection	11.76	59.26	1.28	16	1.12
	4–3	collision type: rear-end, intersection type: T-intersection, ownership: private	11.76	64.00	1.38	16	1.08
	4–4	collision type: rear-end, time: daytime	16.18	57.89	1.25	22	1.10
	4–5	collision type: rear-end, light condition: daylight	16.18	57.89	1.25	22	1.10
	4–6	collision type: rear-end, light condition: daylight, ownership: private	16.18	61.11	1.32	22	1.06
	4–7	collision type: rear-end, day: weekday	16.18	56.41	1.22	22	1.07
	4–8	collision type: rear-end, day: weekday, time: daytime	13.97	63.33	1.37	19	1.12
	4–9	collision type: rear-end, day: weekday, light condition: daylight	13.97	63.33	1.37	19	1.12
	4–10		20.59	56.00	1.21	28	1.06

(*continued on next page*)

Table 4 (*continued*)

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
		collision type: rear-end, ownership: private					
	4–11	collision type: rear-end, ownership: private, time: daytime	16.18	61.11	1.32	22	1.09

consequent of the rule. Using this supervised data mining approach could provide insights into which groups of risk factors lead to each casualty severity level. As noted in the literature, policymakers may then use them by focusing on developing countermeasures to which aim at improving or eliminating such variables (Samerei et al., 2021).

4.3.1. Rules for motorcycle crashes leading to KSI and MINJ (full data)

From Table 3, fourteen strong rules were obtained for KSI crashes, and twenty-one strong rules were obtained for MINJ crashes. Regarding the KSI crashes, it was identified that most rules were associated with collision type, casualty age, type of control at the intersection, and the intersection type. The first rule in this category, Rule 1–1 (collision type: right-angle -> Casualty severity: KSI), indicates that if a right-angle crash occurs at an intersection, it is likely to have very severe consequences as fatal/serious injuries (KSI) is highly likely (lift = 1.25). Besides, according to the rule, right-angle collisions gave rise to a proportion of KSI crashes equal to 67.14% (based on the confidence value). Further, it was identified in Rule 1–2 that the occurrence of right-angled collisions at a four-legged intersection further increased the probability of observing fatal/serious injuries (lift = 1.35). This finding is intuitive as motorcycle riders are more vulnerable to right-angle collisions at intersections due to exposure to red runners/right-of-way rule flouters (Haque et al., 2008); (Haque and Chin, 2010). According to (Talbot et al., 2020); motorcycle crashes at intersections are likely to occur due to overspeeding or missed/late observations on the part of either the rider or driver. This finding is consistent with the literature that identified that motorcycle crash fatality tended to increase in right-angle crashes (Pai and Saleh, 2008); (Rifaat et al., 2012); (Savolainen and Mannering, 2007).

According to Rule 1–3, the risk of fatal/serious injury crashes involving younger casualty's are likely to be high (lift = 1.17), giving rise to a proportion of KSI crashes equal to 62.83%. A combination of items such as median and shoulder presence, right-angled collisions, non-signalized intersections, four-legged intersection type, and two-vehicle involvement produced a further increase in lift values (Rules 1–4 to 1–10). This result shows that the fatality of crashes involving younger casualties at non-signalized four-legged intersections and segments with medians and shoulder presence is likely to increase the chance that higher fatalities are observed. In particular, the rule with the highest lift in this category (Rule 1–7) indicates that two or more vehicle motorcycle-involved crashes at non-signalized intersections are likely to be fatal or severe if the casualty age is less than 30. These findings are in line with the results of studies in the literature. The study conducted by (Abrari Vajari et al., 2020) also identified that motorcycle crashes occurring at non-signalized intersections (uncontrolled/stop or give-way sign) are likely to be fatal. Besides, the study showed that multi-vehicle crashes involving motorcycles were found to be mostly fatal.

According to Rules 1–11 and 1–12, motorcycle crashes at four-legged intersections with shoulders present increase the chance of observing fatal or serious injuries. Interestingly, riders with full license status were identified as the group of motorcycle riders with a high chance of being involved in KSI crashes (Rule 1–13 and 1–14). A likely explanation of these results is that motorcyclists with full licenses are likely to be overconfident as they ride on straight roads (while making going-ahead

Table 5
List of interesting rules for motorcycle crashes at non-signalized intersections

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
Casualty severity: KSI	5–1	collision type: right-angle	19.84	71.43	1.32	25	n.a.
	5–2	collision type: right-angle, median: present	11.90	78.95	1.46	15	1.11
	5–3	collision type: right-angle, median: present, surface condition: good	11.90	83.33	1.54	15	1.06
	5–4	collision type: right-angle, shoulder: none	10.32	76.47	1.42	13	1.07
	5–5	collision type: right-angle, shoulder: none, weather: clear	10.32	81.25	1.51	13	1.06
	5–6	casualty age: less than 30	29.37	68.52	1.27	37	n.a.
	5–7	casualty age: less than 30, collision type: right-angle	11.90	78.95	1.46	15	1.15
	5–8	casualty age: less than 30, median: present	15.87	76.92	1.43	20	1.12
	5–9	casualty age: less than 30, number of vehicles: two-vehicle crash	26.98	73.91	1.37	34	1.08
	5–10	median: present	31.75	59.70	1.11	40	n.a.
	5–11	median: present, time: nighttime	11.90	75.00	1.39	15	1.26
	5–12	median: present, shoulder: none	14.29	75.00	1.39	18	1.26
	5–13	median: present, violation: inattentive	14.29	75.00	1.39	18	1.26
Casualty severity: MINJ	6–1	casualty age: from 30 to 50	26.98	59.65	1.32	34	n.a.
	6–2	casualty age: from 30 to 50, day: weekend	11.11	73.68	1.63	14	1.24
	6–3	casualty age: from 30 to 50, day: weekend, shoulder: present	11.11	82.35	1.82	14	1.12
	6–4	casualty age: from 30 to 50, time: nighttime	11.11	70.00	1.55	14	1.17
	6–5	casualty age: from 30 to 50, collision type: rear-end	10.32	68.42	1.51	13	1.15
	6–6	casualty age: from 30 to 50, collision type: rear-end, shoulder: present	10.32	92.86	2.05	13	1.36
	6–7	casualty age: from 30 to 50, collision type: rear-end, license status: unknown	10.32	72.22	1.60	13	1.06
	6–8	casualty age: from 30 to 50, shoulder: present	20.63	68.42	1.51	26	1.15
	6–9	casualty age: from 30 to 50,	26.19	64.71	1.43	33	1.08

(continued on next page)

Table 5 (*continued*)

Consequent	ID	Antecedent	S%	C%	Lift	N	LIC
		license status: unknown					
	6–10	casualty age: from 30 to 50, license status: unknown, shoulder: present	19.84	73.53	1.63	25	1.14
	6–11	casualty age: from 30 to 50, license status: unknown, time: nighttime	10.32	72.22	1.60	13	1.12
	6–12	casualty age: from 30 to 50, license status: unknown, median: none	10.32	72.22	1.60	13	1.12
	6–13	casualty age: from 30 to 50, license status: unknown, maneuver: going-ahead	26.19	68.75	1.52	33	1.06
	6–14	casualty age: from 30 to 50, median: none	11.11	63.64	1.41	14	1.07
	6–15	casualty age: from 30 to 50, intersection type: T-intersection	19.05	63.16	1.40	24	1.06
	6–16	casualty age: from 30 to 50, intersection type: T-intersection, shoulder: present	15.87	74.07	1.64	20	1.17
	6–17	casualty age: from 30 to 50, intersection type: T-intersection, maneuver: going-ahead	19.05	68.57	1.52	24	1.09
	6–18	casualty age: from 30 to 50, intersection type: T-intersection, light condition: daylight	15.08	67.86	1.50	19	1.07
	6–19	casualty age: from 30 to 50, maneuver: going-ahead	26.98	62.96	1.39	34	1.06
	6–20	casualty age: from 30 to 50, maneuver: going-ahead, shoulder: present	20.63	74.29	1.64	26	1.18
	6–21	casualty age: from 30 to 50, maneuver: going-ahead, time: nighttime	11.11	70.00	1.55	14	1.11
	6–22	casualty age: from 30 to 50, maneuver: going-ahead, median: none	11.11	66.67	1.47	14	1.06
	6–23	day: weekend	15.87	57.14	1.26	20	n.a.
	6–24	day: weekend, shoulder: present	15.08	73.08	1.62	19	1.28
	6–25	day: weekend, intersection type: T-intersection	12.70	69.57	1.54	16	1.22
	6–26	day: weekend, intersection type: T-intersection, shoulder: present	12.70	80.00	1.77	16	1.15
	6–27	day: weekend, intersection type: T-intersection, number of vehicles: two-vehicle crash	11.11	77.78	1.72	14	1.12
	6–28	day: weekend, intersection type: T-intersection, license status: unknown	11.11	73.68	1.63	14	1.06
	6–29	day: weekend, ownership: private	15.08	63.33	1.40	19	1.11
	6–30	day: weekend, ownership: private, shoulder: present	15.08	73.08	1.62	19	1.15
	6–31	median: none	23.81	51.72	1.14	30	n.a.
	6–32	median: none, time: nighttime	11.90	78.95	1.75	15	1.53
	6–33	median: none, time: nighttime, weather: clear	10.32	86.67	1.92	13	1.10
	6–34	shoulder: present	33.33	50.60	1.12	42	n.a.
	6–35	shoulder: present, time: nighttime	11.90	65.22	1.44	15	1.29
	6–36	shoulder: present, violation: none	16.67	55.26	1.22	21	1.09
	6–37	collision type: rear-end	13.49	50.00	1.11	17	n.a.
	6–38	collision type: rear-end, shoulder: present	12.70	66.67	1.47	16	1.33

maneuvers). This could lead them to engage in risky behaviors as such roads give them an elevated sense of safety. They are likely to adjust to the road condition by speeding, which, in tandem with the literature (Pai and Saleh, 2007), can cause catastrophic consequences, especially at non-signalized intersections. Besides, as previously explained, many of the full license holders did not acquire them through legitimate means (Obiri-Yeboah et al., 2021). Thus, they may not have acquired the appropriate training to ride on the roadways.

Regarding the result concerning shoulder presence, most riders in Ghana utilize the road shoulders at intersections as a bypass in order to move to the front of the vehicle queue during the red phase. A similar observation where motorcyclists use other lanes to move to the front of the vehicle queue during the red phase was reported in a study investigating the behavior of motorcyclists in Singapore (Haque et al., 2008). This behavior could lead to crashes involving vehicles or pedestrians using the roadway shoulders due to the lack of sidewalks (Aidoo and Amoh-Gyimah, 2020). In line with previous studies, it was identified that shoulder absence reduced the probability of fatal injuries and increased the probability of PDO crashes (Wahab and Jiang, 2019). Another study discovered that shoulder width plays a key role in increasing the severity of motorcycle crashes (Schneider and Savolainen, 2011). The rules regarding shoulder presence, four-legged intersections, and full license statuses have high lift values, indicating strong associations with fatality. Thus, it would be important for authorities to pay much attention to the design of roads, particularly shoulders at intersections.

Regarding MINJ crashes, several interesting observations were made. The most common crash-contributory factors in this category correspond to crash type (rear-end), intersection type (T-intersection), day of the week (weekend), license status (unknown), and shoulder (present). First, crashes involving people in the mid-age range (30 s to 50 s) are likely to have minor severities (Rule 2–1). In particular, the item, {casualty age: from 30 to 50} exhibited a lift of 1.20 and gave rise to a proportion of MINJ crashes equal to 55.17%. Combining this item with the unknown license status variable yielded the highest lift rule in this category (Rule 2–2; lift = 1.42). Again, the lift value of the parent rule (Rule 2–1) also increases to 1.30 if the crash occurred in a non-signalized intersection, signifying a high association of non-signalized intersections and mid-aged casualties with MINJ crashes (Rule 2–5). Essentially, the rule with the second-highest lift value in the MINJ category shows very strong association between crashes involving people in the mid-age range, unknown license status, non-signalized intersection types, and minor injury outcomes (Rule 2–6). Besides, MINJ crashes involving this age cohort are likely to occur in the nighttime and at road segments with shoulder presence (Rules 2–3, 2–4, and 2–8).

According to Rule 2–9, rear-end crashes are likely to be MINJ. Besides, rear-end cashes occurring at T-intersections with shoulder presence during the daytime/daylight conditions are likely to be MINJ. Crashes at T-intersections were also found to be a key contributor to MINJ crashes Rule 2–14. In the literature, a study conducted in the UK showed that motorcyclist injury severity is likely to be less severe in rear-end crashes during the daytime (Pai and Saleh, 2008). This finding is intuitive as the increased traffic volumes during the daytime is likely to cause motorists to reduce their speeds.

According to Rules 2–15 and 2–16, crashes at T-intersections are likely to be MINJ when the crash occurs at a shoulder-present area or the motorcycle rider commits no violations. Besides, crashes occurring during the weekend irrespective of the intersection type, involving private motorcycles, and at shoulder-present intersections are identified to have strong correlations with MINJ crashes. The finding regarding the reduced severity of motorcycle rider not-at-fault crashes is counterintuitive, as a previous study in Vietnam identified that second-party motorcycle crashes present more lethal consequences than first-party crashes (Nguyen et al., 2021). In the case of Ghana, over 50% of motorcycle riders and their passengers do not wear protective gear such as helmets which could help protect them against extremities (Nimako Aidoo et al., 2018). As such, they are likely to be injured in crashes. Nevertheless, the findings of this study may be reasonable since, whereas the study from Vietnam considered crashes at all locations, our result applied to crashes at T-intersections.

It is worth pointing out that, whereas this study identified that motorcycle crashes at intersections on weekends were highly associated with minor injuries, some studies from developed countries showed otherwise. An Australian-based study suggests that crashes on weekends are likely to be fatal. According to the authors, this finding might be the case as many riders are likely to speed during the weekends (Abrari Vajari et al., 2020). However, according to a UK-based study, weekday crashes positively influence the severity at signalized intersections and negatively influence crash severity at non-signalized three-legged intersections (Pai and Saleh, 2007). These contrasting findings highlight the need to explore the chains of risk factors influencing severity at signalized and non-signalized intersections separately.

4.3.2. Rules for motorcycle crashes at signalized intersections leading to KSI and MINJ

Association rules for motorcycle crashes at signalized intersections are presented in Table 4. Interestingly, the majority of the rules resulting in KSI involved motorcycle riders with full license status. Other common variables include daylight condition, right-angle, and shoulder and median presence. The first rule in this category shows that, for right-angled crashes at signalized intersections, the proportion of KSI was 62.86%. Besides, if a right-angled motorcycle crash occurs at a signalized intersection, the chance of fatal/serious injuries is highly likely to occur (Rule 3–1; lift = 1.19). Additionally, according to Rule 3–2, the lift value further increases to 1.32 when the right-angled crash occurred on a weekday, indicating a much stronger association between right-angle crashes and weekend crashes with KSI crashes. These findings were in line with the literature that identified that angle motorcycle crashes at signalized intersections were associated with higher levels of fatalities (Pai and Saleh, 2007).

Rules 3–3 to 3–10 show strong associations between KSI and variables, namely full license status, no violations, daytime, daylight condition, median presence, shoulder presence, two-vehicle crashes, and going-ahead maneuvers. In particular, motorcyclists with full license status had an increased chance of KSI crashes (Rule 3–3), and a combination of non-violation of traffic regulations and daytime variables increased the probability of KSI (Rules 3–4 and 3–5). Rule 3–7, which has the highest lift value of 1.46 in this category, shows that riders with full license status who ride during daylight conditions are likely to be involved in fatal crashes at signalized intersections as they partake in "going-ahead" maneuvers. It is commonplace to find motorcyclists using flush medians (unraised) and shoulders as they weave through traffic or at intersections. According to the literature, weaving at the area near intersections exposes motorcyclists to severe crashes with unsuspecting road users. Besides, these riders often utilize the early portion of the green phase, exposing them to right-angled crashes with red-light runners (Aidoo and Amoh-Gyimah, 2020); (Haque et al., 2008); (Haque and Chin, 2010). Besides, regarding temporal features, the findings are plausible and in line with previous studies showing that motorcycle crashes occurring during the daytime or peak time are likely to be severe relative to the nighttime (Rifaat et al., 2012); (Wahab and Jiang, 2019).

The next rule in this category with the highest lift (Rule 3–9; lift = 1.40) shows that crashes involving motorcycle riders with full license status who are involved in a two-vehicle involved crash at median-present road segments with signalized intersections are likely to yield KSI's. Besides, Rule 3–10 shows that shoulder presence is also a critical contributory factor of KSI motorcycle crashes at signalized intersections involving riders with full license status. The finding that shoulder presence is associated with KSI crashes is intuitive, as motorcyclists in Ghana have the habit of riding on the road's shoulders, although the activity is illegal. A recent study from Ghana shows that road shoulders are undoubtedly critical locations where motorcycle crashes occur (Jack et al., 2021). Motorcycle riders also have a high risk of having critical interactions with pedestrians or other hazardous objects placed around the roadway's shoulders in the event of a run-off-road crash or as they dodge traffic while using the road shoulders. Consistent with our finding, the presence of roadway shoulder is likely to cause fatal crashes, especially when the pavement condition is poor (Wahab and Jiang, 2019).

Further, the rules leading to MINJ crashes at signalized intersections were investigated. Interestingly, all the rules in this category contain the variable rear-end. Other items identified to be associated with KSI crashes at signalized intersections include variables, namely private ownership, daylight, daytime, weekday, and T-intersection. The first rule in this category depicts that 21.59% of the motorcycle crashes at signalized intersections were rear-end crashes. The high lift value of 1.14 shows a strong correlation between rear-end crashes and MINJ crashes at signalized intersections. Due to the static nature of at least one involved vehicle, rear-end crashes are often less severe than right-angle or turn-related crashes. Combining this rule with both private ownership and T-intersections yielded a rule with the highest lift value in this category (Rule 4–3, lift = 1.38). Moreover, weekday crashes occurring during the daytime or during daylight conditions are likely to result in MINJ (Rules 4–8 and 4–9). Weekdays crashes at signalized intersections are associated with less severe outcomes due to the lower operating speeds from congestion and delays due to the traffic signals.

4.3.3. Rules for motorcycle crashes at non-signalized intersections leading to KSI and MINJ

Rules for motorcycle crashes at non-signalized intersections were also mined, with both KSI and MINJ as consequents (see Table 5. Variables, namely right-angle, casualty age less than 30, median presence, and shoulder absence, were identified as common contributory factors associated with KSI motorcycle crashes at non-signalized intersections. Rule 5–1 shows that right-angled crashes at non-signalized intersections are highly likely to be KSI crashes – consistent with the literature (Beeharry et al., 2020); (Rifaat et al., 2012). The rule also shows that, for right-angled crashes at non-signalized intersections, the proportion of fatal/serious injury casualty outcomes is 71.43%. According to Rule 5–2, if the right-angled motorcycle crash occurred at a non-signalized intersection with median-present, the probability of fatality increases by 1.46 times. Again, the probability of a KSI crash increases (lift = 1.42) at non-signalized intersections if the road surface condition is good. Right-angled crashes occurring during clear weather periods and at locations without shoulders are likely to increase the probability of fatality by 1.51 (the highest lift value in the category). These findings are in line with previous studies that identified that good roadway condition and weather are critical determinants of increased injury severity in both motorcycle and general vehicle crashes (Tamakloe et al., 2021; Tamakloe et al., 2020); (Wahab and Jiang, 2019), and is plausible as the increased visibility and comfort on cruising on good roads could make the driver/rider less careful and/or speed and thus increasing the probability of fatal/serious injuries (as previously explained). Besides, the absence of shoulders suggests that riders cannot maneuver to safety when faced with a potential crash (Aidoo and Amoh-Gyimah, 2020).

Rules 5–7 to 5–11 suggest that young road users are likely to sustain severe or fatal injuries in motorcycle crashes at non-signalized intersections. Besides, crash fatality is likely to increase in right-angled collisions when two vehicles are involved in the crash and when there is a median present. As depicted by Rule 5–11 (lift = 1.11), median presence and nighttime variables strongly influence KSI crashes. Similarly, a Mauritius-based study also identified that the nighttime variable was associated with severe PTW crashes (Beeharry et al., 2020). Also, if a motorcycle crash at the median-present segment occurs during the nighttime, due to inattention, or at a segment with no shoulder, the consequences are likely to be fatal (Rules 5–12 and 5–13). In line with the study conducted by (Talbot et al., 2020); missed observation owing to attention allocation issues and priority errors is likely to lead to fatal motorcycle crashes at intersections. These findings are plausible since being less careful while driving or riding could cause motorists to make deadly mistakes on the roadway. Moreover, riding during the night where traffic volumes are low (at median-separated segments) tends to induce the rider to overspeed as the motorist perceives the conditions to be safer, leading to increased frequency and probability of fatal/serious injuries (Hong et al., 2020).

Using the apriori algorithm for mining rules of crashes leading to minor injuries at non-signalized intersections, a total of 38 rules were identified. More than half of these rules had the item {casualty age: from 30 to 50} as an antecedent. In particular, the first rule in that category signifies that people in their mid-ages are likely to sustain minor injuries in motorcycle crashes at non-signalized intersections. This rule had a high lift value of 1.32, indicating that crashes related to people in this age group did not happen by chance. Notably, the results in the preceding rules showed strong associations with factors such as rear-end, unknown license status, and going-ahead maneuvers.

The rule with the highest lift value (Rule 6–6, lift = 2.05) shows that 10.32% of motorcycle crashes at non-signalized intersections were rear-ended MINJ crashes involving road users in their mid-ages (30 s to 50 s) at segments with shoulders present. The rule also shows that 92.86% of the motorcycle crashes which resulted in MINJ crashes were rear-ended collisions at a shoulder-present segment involving mid-aged riders. The high lift of this rule shows a strong correlation between MINJ crashes and the road users in the mid-aged cohort, shoulder presence, and rear-ended crashes. A similar finding is presented by Rules 6–37 and 6–38. It is noteworthy that the finding that rear-end motorcycle crashes at non-signalized intersections have strong associations with minor injuries in Ghana is consistent with the UK-based study that identified that motorcycle crashes at uncontrolled/stop/yield intersections are less likely to be severe (Pai and Saleh, 2007) However, it is inconsistent with the finding reported in the study that explored PTW crashes at non-signalized intersections in Mauritius. In the literature, it has been documented that rear-end crashes at signalized intersections are common when the speed limits of the road are high or when the leading vehicle makes an abrupt stop unknowing to the vehicle following when the yellow traffic light is displayed (Haque et al., 2010). At unsignalized intersections, these crashes may also happen when the leading vehicle makes a sudden decision to yield to other vehicles, but the following vehicle is unable to stop. Due to the reduced speeds at unsignalized intersections and high congestion levels on most roads in most Ghanaian cities, including the GAR (Agyapong and Ojo, 2018); (Møller-Jensen, 2021), it is likely that the rear impact would not result in fatalities.

Other rules in this category show a high correlation between weekend crashes, median absence, nighttime crashes, and clear weather with MINJ crashes (Rules 6–23 to 6–35). In particular, Rule 6–33 shows that crashes occurring at median-absent and non-signalized intersections during the clear weather at nighttime are likely to be less severe. This finding may be the case as motorists are more careful due to the median absence and reduced visibility during the night.

4.4. Comparison of results from the binary logit regression model and association rule mining technique

Exploring the crash data using both the BLR model and the ARM approach led to the identification of the impact of individual risk factors on motorcycle crash casualty severity outcomes and the combination of risk factors that collectively lead to each severity outcome. As presented in Table 6, some statistically significant variables in the BLR model formed part of the rules leading to specific crash severity outcomes obtained using the ARM approach. It is noteworthy that, for the signalized intersection data, the variables identified in the rules were not significant in the BLR model. This could be due to the selection of thresholds used for the ARM analysis, which led to a reduced number of strong rules. However, both model results are comparable.

Regarding the non-signalized intersection data, the BLR model showed that right-angle collisions were likely to result in KSI casualty severity. The ARM approach further showed that a combination of factors such as median presence, good road conditions, clear weather, shoulder absence, and right-angled collision type of crashes are highly likely to result in KSI casualty severity outcome in case of motorcycle crashes (rules 5–1 to 5–5). Again, in both the total data and the non-signalized intersection data, it was observed from the BLR model that young motorcycle crash casualties are likely to be killed or sustain serious injuries. Besides, from the ARM results, these young people are likely to have KSI severity outcomes if the motorcycle crash involving another vehicle is a right-angle collision and occurs at a section with median presence. The four-legged intersection type was identified to be highly linked with these crashes in the full/total data category. Concerning the crashes resulting in MINJ severity outcomes, unknown license status was found to be significant in the BLR model in both models using the Total data and non-signalized intersection data. The ARM technique showed that these non-signalized intersection motorcycle-

Table 6
Comparing results from both binary logit model and association rules mining.

Location of crash	BLR		ARM	
	Casualty severity outcome	Significant variable	Rules	Variables contained in the rules
Total	KSI	Casualty age: less than 30	1–3 to 1–11	Casualty age: less than 30, control: non-signalized, number of vehicles: two-vehicle crash, shoulder: present, median: present, intersection type: four-legged, collision type: right-angle
Total	MINJ	Intersection type: T-intersection	2–11, 2–14 to 2–16	Collision type: rear-end, intersection type: T-intersection, shoulder: present, violation: none
Total	MINJ	License status: unknown	2–2 to 2–3, 2–6, 2–18 to 2–19	Casualty age: from 30 to 50, license status: unknown, shoulder: present, control: non-signalized, day: weekend, ownership: private
NSIG	KSI	Collision type: right-angle	5–1 to 5–5	Collision type: right-angle, median: present, surface condition: good, shoulder: none, weather: clear
NSIG	KSI	Casualty age: less than 30	5–6 to 5–9	Casualty age: less than 30, collision type: right-angle, median: present, number of vehicles: two-vehicle crash
NSIG	MINJ	Median: none	6–31 to 6–33	Median: none, time: nighttime, weather: clear
NSIG	MINJ	License status: unknown	6–7, 6–9 to 6–13	Casualty age: from 30 to 50, license status: unknown, shoulder: present, time: nighttime, median: none, maneuver: going-ahead, collision type: rear-end

Note: Total: motorcycle-involved crash at intersections (signalized/non-signalized); NSIG: motorcycle-involved crash at a non-signalized intersection; BLR: binary logit regression; ARM: association rules mining

involved crashes were strongly associated with young crash casualties and shoulder presence. Interestingly, nighttime was highly associated with MINJ severity outcome crashes at non-signalized intersections.

Overall, the findings show that the results from both models are comparable and together present valuable insights. Based on the results, we can diagnose the motorcycle safety problem as being caused by a chain of factors, from which researchers and safety experts should consider employing for crash severity mitigation in Ghana.

5. Conclusion and policy recommendations

Investigating the chains of factors influencing motorcycle crashes in developing countries has become crucial due to the high number of casualties associated with this vehicle type. Given that the use of motorcycles in these nations continues to increase, it is essential to conduct more detailed studies to find targeted countermeasures to help reduce the level of carnage on the roadways.

The objective of this study was to explore and identify hidden chains of factors among the contributory factors in a motorcycle crash database leading to both casualty KSI and MINJ crashes at signalized and non-signalized intersections in Ghana using a plausible data mining approach. Intersection-related motorcycle crash data were segregated based on the traffic control present. This study identified interesting latent chains of factors influencing casualty injury severity at both signalized and non-signalized intersections by applying the association rule mining technique. As a benchmark, the BLR model was also used for the investigation conducted in this study. The results from both models were identified to agree with each other. As previously discussed, the traditional BLR model only provided insights into the impact of individual variables on casualty severity outcomes, whereas the ARM discovered hidden rules consisting of chains of factors influencing each casualty severity outcome. The findings from the ARM provide an added advantage in that it provides the chain of factors that collectively influence motorcycle crash casualty severity outcomes. Thus, the severity of these crashes could be easily reduced by eliminating or reducing some or all of these factors identified in the factor chains.

Among the rules obtained, some of the key findings are as follows:

- Regarding crashes at signalized intersections, both median and shoulder presence were associated with KSI crashes. Moreover, while KSI crashes at non-signalized intersections are associated with median presence and shoulder absence, MINJ crashes are associated with median absence and shoulder presence.
- Right-angle crashes led to KSI, and rear-end crashes led to MINJ irrespective of the control type at the intersection.
- Casualty age was more prominent in rules belonging to the non-signalized crash model. While younger casualties sustained fatal/serious injuries, those in the mid-ages had minor injuries.
- While weekday crashes were more correlated with crashes at signalized intersections, weekend crashes were found to be more associated with crashes at non-signalized intersections.
- The variable for daytime was more associated with rules for crashes at signalized intersections. On the other hand, the variable for nighttime was more associated with crashes occurring at non-signalized intersections. However, this finding needs to be carefully interpreted. For example, higher daytime crashes may be associated with high exposure to motorcycle travels during the daytime.

The findings suggest that the inefficient licensing system and the activities of the middlemen at DVLA offices, which leads to the acquisition of full licenses without thorough in-person screening, is a major reason why full license status holders in the GAR are likely to be involved in KSI crashes. Besides, the nature of the roads in the GAR, the indiscipline behavior of motorcycle riders, and the operation style of motorcycle riders (particularly "okada"), among others, were identified as critical reasons for the analysis results. Based on the findings, we suggest policies for addressing the fatalities associated with motorcycle crashes both locally and internationally.

- First, while increased enforcement is needed to ensure that motorcycle riders observe traffic regulations and wear protective gear, it is also crucial to increase rider education, removing the obstacles preventing people from acquiring licenses themselves (including the strict bureaucracy and the middlemen or "Goro boys" operating at the local DVLA offices), and revisiting/reforming the training given to both riders and drivers before issuing licenses. Re-education of current full license holders could be achieved through the media or campaigns targeting younger people in schools, particularly how to use non-signalized intersections safely..
- Currently, Ghana has not adequately enforced vehicle standards policies (WHO, 2018). Collisions such as rear-ends could be minimized by enacting and enforcing policies regarding adopting

appropriate state-of-the-art vehicle standards such as motorcycle anti-lock braking systems.

- To reduce the occurrence and severity of right-angle collisions, it would be worthwhile to provide red-light cameras at intersections. These devices are noted as being particularly useful in reducing the right-angle vulnerability of motorcyclists (Haque and Chin, 2010).
- To reduce the severity of crashes at shoulder/median-present intersection segments, it would be worthwhile to enforce regulations banning motorcyclists from using those areas strictly. Besides, providing motorcyclist protection devices such as roller barriers instead of sharp guardrails at roadway shoulders in crash hotspots could reduce the severity of single-vehicle crashes. Further, it is imperative to allocate enough budget for maintaining overgrown shoulders and medians as they are identified to exacerbate the lack of visibility at intersections (Aidoo and Amoh-Gyimah, 2020).
- The use of retroreflective markings, provision of warning signs and reduced speed limit posts at curvatures, and the installation of other safety devices such as reflectors and rumble strips at the edges and the centerline of median-absent and shoulder-absent roads could help keep driver's alert.
- Inobservance of traffic regulations is common among motorcycle riders in Ghana (Jack et al., 2021). It is imperative to increase surveillance on roadways, particularly at signalized intersections on weekdays during the daytime and at non-signalized intersections during the night on weekends, to check risky and inattentive riding at intersections. This could be achieved through physical policing and the deployment of traffic cameras.
- In the long run, it is imperative to consider the provision of funds for the redesign of roadways and the incorporation of motorcycle lanes, safer intersections, or wider shoulders where necessary.
- It is recommended to critically consider restricting or regulating the use of motorcycles as commercial public transport. Legalizing this business would require the operators to follow strict regulations, reducing crash severity on the roads.

This study contributes to the literature by providing a deeper understanding of the factors influencing motorcycle safety at both signalized and non-signalized intersections. While the issues impacting motorcycle safety in developing countries differ from one another, the findings and recommendations from this study could be used as a basis for the improvement of motorcycle safety in countries worldwide. It is noteworthy that the data used for this study is police-reported, which may be subject to underreporting of motorcycle crashes. Besides, several key rider-level factors such as license status were unknown. It is also noteworthy that the scope of the data used for this study is small (crashes in the capital region of Ghana, GAR, spanning 2016–2018), resulting in a small sample size. Besides, the data did not contain detailed information such as the type of vehicle involved in the motorcycle crash. Thus, insights regarding this variable could not be drawn. In the future, it would be imperative to conduct a nationwide study using a large-sized and more comprehensive dataset to derive valuable results for improving motorcycle safety at intersections.

CRediT authorship contribution statement

Reuben Tamakloe: Conceptualization, Methodology, Software, Data curation, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Subasish Das:** Writing – review & editing ,Validation. **Eric Nimako Aidoo:** Writing – review & editing, Data curation. **Dongjoo Park:** Supervision, Funding acquisition, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the 2021 Research Fund of the University of Seoul (No. 202104241074).

References

Abrari Vajari, M., Aghabayk, K., Sadeghian, M., Shiwakoti, N., 2020. A multinomial logit model of motorcycle crash severity at Australian intersections. J. Safety Res. 73, 17–24. https://doi.org/10.1016/j.jsr.2020.02.008.

Agrawal, R., Imieliński, T., Swami, A., 1993. Mining Association Rules Between Sets of Items in Large Databases. ACM SIGMOD Rec. DOI 10 (1145/170036), 170072.

Agyapong, F., Ojo, T.K., 2018. Managing traffic congestion in the Accra Central Market. Ghana. J. Urban Manag. 7 (2), 85–96. https://doi.org/10.1016/j.jum.2018.04.002.

Aidoo, E.N., Amoh-Gyimah, R., 2020. Modelling the risk factors for injury severity in motorcycle users in Ghana. J. Public Health (Bangkok) 28 (2), 199–209. https://doi.org/10.1007/s10389-019-01047-7.

Albalate, D., Fernández-Villadangos, L., 2010. Motorcycle Injury Severity in Barcelona: The Role of Vehicle Type and Congestion. Traffic Inj. Prev. 11 (6), 623–631. https://doi.org/10.1080/15389588.2010.506932.

Amoh-Gyimah, R., Aidoo, E.N., Akaateba, M.A., Appiah, S.K., 2017. The effect of natural and built environmental characteristics on pedestrian-vehicle crash severity in Ghana. Int. J. Inj. Contr. Saf. Promot. 24 (4), 459–468. https://doi.org/10.1080/17457300.2016.1232274.

Appiah, A.N., 2018. The Perception of Risk and Its Implications on the Operations of Motor Taxi (Okada). in the Greater Accra Region. University of Ghana.

Armah, F., Yawson, D., Pappoe, A.A.N.M., 2010. A Systems Dynamics Approach to Explore Traffic Congestion and Air Pollution Link in the City of Accra. Ghana. Sustainability 2 (1), 252–265. https://doi.org/10.3390/su2010252.

Beeharry, R., Goodary, R., Ratanavaraha, V., 2020. Impact of road infrastructure and environment on powered-two-wheelers injury severity at non-signalized intersections in a developing country – Mauritius. J. Emerg. Trends Eng. Appl. Sci. 11 (6), 219–228.

Boateng, F.G., 2021. Why Africa cannot prosecute (or even educate) its way out of road accidents: insights from Ghana. Humanit. Soc. Sci. Commun. 8 (1), 13. https://doi.org/10.1057/s41599-020-00695-5.

Brin, S.C., Motwani, R., 1998. Beyond Market Baskets: Generalizing Association Rules to Dependence Rules. Data Min. Knowl. Discov. 2, 39–68. https://doi.org/10.1023/A:1009713703947.

Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., Jalayer, M., 2019. Supervised association rules mining on pedestrian crashes in urban areas: identifying patterns for appropriate countermeasures. Int. J. Urban Sci. 23 (1), 30–48. https://doi.org/10.1080/12265934.2018.1431146.

Das, S., Kong, X., Tsapakis, I., 2019. Hit and run crash analysis using association rules mining. J. Transp. Saf. Secur. 1–20 https://doi.org/10.1080/19439962.2019.1611682.

Das, S., Sun, X., Goel, S., Sun, M., Rahman, A., Dutta, A., 2020. Flooding related traffic crashes: findings from association rules. J. Transp. Saf. Secur. 1–19 https://doi.org/10.1080/19439962.2020.1734130.

Das, S., Tamakloe, R., Zubaidi, H., Obaid, I., Alnedawi, A., 2021. Fatal pedestrian crashes at intersections: Trend mining using association rules. Accid. Anal. Prev. 106306 https://doi.org/10.1016/j.aap.2021.106306.

Haque, M.M., Chin, H.C., 2010. Right-Angle Crash Vulnerability of Motorcycles at Signalized Intersections. Transp. Res. Rec. J. Transp. Res. Board 2194 1, 82–90. https://doi.org/10.3141/2194-10.

Haque, M.M., Chin, H.C., Huang, H., 2010. Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections. Accid. Anal. Prev. 42 (1), 203–212. https://doi.org/10.1016/j.aap.2009.07.022.

Haque, M.M., Chin, H.C., Huang, H., 2008. Examining Exposure of Motorcycles at Signalized Intersections. Transp. Res. Rec. J. Transp. Res. Board 2048 1 , 60–65. 10.3141/2048-08.

Hong, J., Tamakloe, R., Park, D., 2020. Discovering Insightful Rules among Truck Crash Characteristics using Apriori Algorithm. J. Adv. Transp. 2020, 1–16. https://doi.org/10.1155/2020/4323816.

Hong, J., Tamakloe, R., Park, D., 2020. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. Accid. Anal. Prev. 142, 105497 https://doi.org/10.1016/j.aap.2020.105497.

Jack, J.K.A., Amoah, E.K.S., Hope, E., Okyere, F., 2021. Should Ghana Legalize the Commercial Use of Motor Bikes and Tricycles as Means of Public Transport? A Case Study of Five Selected Regions in Ghana. J. Econ. Bus. 4 1 . 10.31014/aior.1992.04.01.333.

Jiang, F., Yuen, K.K.R., Lee, E.W.M., 2020. Analysis of motorcycle accidents using association rule mining-based framework with parameter optimization and GIS technology. J. Safety Res. 75, 292–309. https://doi.org/10.1016/j.jsr.2020.09.004.

Kim, J.-K., Kim, S., Ulfarsson, G.F., Porrello, L.A., 2007. Bicyclist injury severities in bicycle–motor vehicle accidents. Accid. Anal. Prev. 39 (2), 238–251. https://doi.org/10.1016/j.aap.2006.07.002.

Li, X., Wang, Y., 2020. Scene Identification of Single-Motorcycle Death Accidents Based on Binary Logistic Regression and Association Rules, in: CICTP 2020. American Society of Civil Engineers, Reston, VA, pp. 4573–4584. 10.1061/9780784483053.380.

Lin, M.-R., Huang, W., Hwang, H.-F., Wu, H.-D.-I., Yen, L.-L., 2004. The effect of crash experience on changes in risk taking among urban and rural young people. Accid. Anal. Prev. 36 (2), 213–222. https://doi.org/10.1016/S0001-4575(02)00150-1.

Manski, C.F., McFadden, D., 1981. A structural analysis of discrete data with econometric applications. MIT Press, Cambridge, MA.

Møller-Jensen, M., 2021. Frictions of everyday mobility: traffic, transport and gendered confrontations on the roads of Accra. Mobilities 16 (4), 461–475. https://doi.org/10.1080/17450101.2021.1917969.

Montella, A., 2011. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. Accid. Anal. Prev. 43 (4), 1451–1463. https://doi.org/10.1016/j.aap.2011.02.023.

Montella, A., Aria, M., D'Ambrosio, A., Mauriello, F., 2012. Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. Accid. Anal. Prev. 49, 58–72. https://doi.org/10.1016/j.aap.2011.04.025.

Montella, A., Aria, M., D'Ambrosio, A., Mauriello, F., 2011. Data-Mining Techniques for Exploratory Analysis of Pedestrian Crashes. Transp. Res. Rec. J. Transp. Res. Board 2237 1, 107–116. https://doi.org/10.3141/2237-12.

Montella, A., de Oña, R., Mauriello, F., Rella Riccardi, M., Silvestro, G., 2020. A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. Accid. Anal. Prev. 134, 105251 https://doi.org/10.1016/j.aap.2019.07.027.

Montella, A., Mauriello, F., Pernetti, M., Rella Riccardi, M., 2021. Rule discovery to identify patterns contributing to overrepresentation and severity of run-off-the-road crashes. Accid. Anal. Prev. 155, 106119 https://doi.org/10.1016/j.aap.2021.106119.

MOT, 2020. National Transport Policy. Accra, Ghana.

Nasri, M., Aghabayk, K., 2021. Assessing risk factors associated with urban transit bus involved accident severity: a case study of a Middle East country. Int. J. Crashworthiness 26 (4), 413–423. https://doi.org/10.1080/13588265.2020.1718465.

Nguyen, D.V.M., Vu, A.T., Polders, E., Ross, V., Brijs, T., Wets, G., Brijs, K., 2021. Modeling the injury severity of small-displacement motorcycle crashes in Hanoi City. Vietnam. Saf. Sci. 142, 105371 https://doi.org/10.1016/j.ssci.2021.105371.

Nimako Aidoo, E., Bawa, S., Amoako-Yirenkyi, C., 2018. Prevalence rate of helmet use among motorcycle riders in Kumasi. Ghana. Traffic Inj. Prev. 19 (8), 856–859. https://doi.org/10.1080/15389588.2018.1509072.

NRSC, 2016. Road Traffic Crash Statistics, 2016. Accra.

Obiri-Yeboah, A.A., Ribeiro, J.F.X., Asante, L.A., Sarpong, A.A., Pappoe, B., 2021. The new players in Africa's public transportation sector: Characterization of auto-rickshaw operators in Kumasi. Ghana. Case Stud. Transp. Policy 9 (1), 324–335. https://doi.org/10.1016/j.cstp.2021.01.010.

Pai, C.-W., 2009. Motorcyclist injury severity in angle crashes at T-junctions: Identifying significant factors and analysing what made motorists fail to yield to motorcycles. Saf. Sci. 47 (8), 1097–1106. https://doi.org/10.1016/j.ssci.2008.12.007.

Pai, C.-W., Saleh, W., 2008. Modelling motorcyclist injury severity by various crash types at T-junctions in the UK. Saf. Sci. 46 (8), 1234–1247. https://doi.org/10.1016/j.ssci.2007.07.005.

Pai, C.-W., Saleh, W., 2007. An analysis of motorcyclist injury severity under various traffic control measures at three-legged junctions in the UK. Saf. Sci. 45 (8), 832–847. https://doi.org/10.1016/j.ssci.2006.08.021.

Pande, A., Abdel-Aty, M., 2009. Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. Saf. Sci. https://doi.org/10.1016/j.ssci.2007.12.001.

Pervez, A., Lee, J., Huang, H., 2021. Identifying Factors Contributing to the Motorcycle Crash Severity in Pakistan. J. Adv. Transp. 2021, 1–10. https://doi.org/10.1155/2021/6636130.

R Core Team, 2017. A Language and Environment for Statistical Computing [WWW Document]. accessed 5.4.20. https://www.r-project.org/.

Rahman, M.H., Zafri, N.M., Akter, T., Pervaz, S., 2021. Identification of factors influencing severity of motorcycle crashes in Dhaka, Bangladesh using binary logistic regression model. Int. J. Inj. Contr. Saf. Promot. 28 (2), 141–152. https://doi.org/10.1080/17457300.2021.1878230.

Rifaat, S.M., Tay, R., de Barros, A., 2012. Severity of motorcycle crashes in Calgary. Accid. Anal. Prev. 49, 44–49. https://doi.org/10.1016/j.aap.2011.02.025.

Salum, J.H., Kitali, A.E., Bwire, H., Sando, T., Alluri, P., 2019. Severity of motorcycle crashes in Dar es Salaam. Tanzania. Traffic Inj. Prev. 20 (2), 189–195. https://doi.org/10.1080/15389588.2018.1544706.

Samerei, S.A., Aghabayk, K., Mohammadi, A., Shiwakoti, N., 2021. Data mining approach to model bus crash severity in Australia. J. Safety Res. 76, 73–82. https://doi.org/10.1016/j.jsr.2020.12.004.

Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accid. Anal. Prev. 39 (5), 955–963. https://doi.org/10.1016/j.aap.2006.12.016.

Schneider, W.H., Savolainen, P.T., 2011. Comparison of Severity of Motorcyclist Injury by Crash Types. Transp. Res. Rec. J. Transp. Res. Board 2265 1, 70–80. https://doi.org/10.3141/2265-08.

Sivasankaran, S.K., Rangam, H., Balasubramanian, V., 2021. Investigation of factors contributing to injury severity in single vehicle motorcycle crashes in India. Int. J. Inj. Contr. Saf. Promot. 28 (2), 243–254. https://doi.org/10.1080/17457300.2021.1908367.

Strauss, J., Miranda-Moreno, L.F., Morency, P., 2014. Multimodal injury risk analysis of road users at signalized and non-signalized intersections. Accid. Anal. Prev. 71, 201–209. https://doi.org/10.1016/j.aap.2014.05.015.

Talbot, R., Brown, L., Morris, A., 2020. Why are powered two wheeler riders still fatally injured in road junction crashes? – A causation analysis. J. Safety Res. 75, 196–204. https://doi.org/10.1016/j.jsr.2020.09.009.

Tamakloe, R., Hong, J., Park, D., 2020. A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data. Accid. Anal. Prev. https://doi.org/10.1016/j.aap.2020.105736.

Tamakloe, R., Lim, S., Sam, E.F., Park, S.H., Park, D., 2021. Investigating factors affecting bus/minibus accident severity in a developing country for different subgroup datasets characterised by time, pavement, and light conditions. Accid. Anal. Prev. 159, 106268 https://doi.org/10.1016/j.aap.2021.106268.

Teng, C.M., 2011. Data, Data, Everywhere, in: Philosophy of Statistics. Elsevier, pp. 1099–1117. doi:10.1016/B978-0-444-51862-0.50034-4.

Vlahogianni, E.I., Yannis, G., Golias, J.C., 2012. Overview of critical risk factors in Power-Two-Wheeler safety. Accid. Anal. Prev. 49, 12–22. https://doi.org/10.1016/j.aap.2012.04.009.

Wahab, L., Jiang, H., 2019. A multinomial logit analysis of factors associated with severity of motorcycle crashes in Ghana. Traffic Inj. Prev. 20 (5), 521–527. https://doi.org/10.1080/15389588.2019.1616699.

Wahab, L., Jiang, H., 2019. A comparative study on machine learning based algorithms for prediction of motorcycle crash severity. PLoS One 14 (4), e0214966. https://doi.org/10.1371/journal.pone.0214966.

Weng, J., Zhu, J.Z., Yan, X., Liu, Z., 2016. Investigation of work zone crash casualty patterns using association rules. Accid. Anal. Prev. 92, 43–52. https://doi.org/10.1016/j.aap.2016.03.017.

WHO, 2018a. Global status report on road safety 2018. Geneva.

WHO, 2018b. Deaths on the roads [WWW Document]. URL https://extranet.who.int/roadsafety/death-on-the-roads/#country_or_area/KOR (accessed 6.21.21).

WHO, 2017. Powered two- and three-wheeler safety: a road safety manual for decision-makers and practitioners. Geneva.

Xu, C., Bao, J., Wang, C., Liu, P., 2018. Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. J. Safety Res. 67, 65–75. 10.1016/j.jsr.2018.09.013.

Yannis, G., Golias, J., Papadimitriou, E., 2005. Driver age and vehicle engine size effects on fault and severity in young motorcyclists accidents. Accid. Anal. Prev. 37 (2), 327–333. https://doi.org/10.1016/j.aap.2004.10.003.

Young, R.K., Liesman, J., 2007. Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model. Accid. Anal. Prev. 39 (3), 574–580. https://doi.org/10.1016/j.aap.2006.10.002.