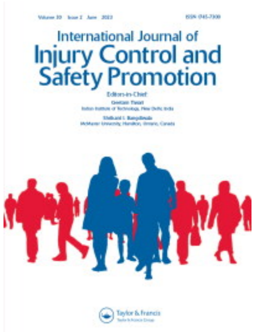

International Journal of Injury Control and Safety Promotion

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/nics20

Using unsupervised learning to investigate injury-associated factors of animal-vehicle crashes

M. Ashifur Rahman, Subasish Das, Xiaoduan Sun, Ming Sun & Md. Mahmud Hossain

To cite this article: M. Ashifur Rahman, Subasish Das, Xiaoduan Sun, Ming Sun & Md. Mahmud Hossain (2023) Using unsupervised learning to investigate injury-associated factors of animal-vehicle crashes, International Journal of Injury Control and Safety Promotion, 30:2, 210-219, DOI: 10.1080/17457300.2022.2125532

To link to this article: https://doi.org/10.1080/17457300.2022.2125532

Published online: 24 Oct 2022.

INTERNATIONAL JOURNAL OF INJURY CONTROL AND SAFETY PROMOTION
2023, VOL. 30, NO. 2, 210–219
https://doi.org/10.1080/17457300.2022.2125532

Using unsupervised learning to investigate injury-associated factors of animal-vehicle crashes

M. Ashifur Rahman[a], Subasish Das[b], Xiaoduan Sun[a], Ming Sun[c] and Md. Mahmud Hossain[d]

[a]Department of Civil Engineering, University of Louisiana, Lafayette, Louisiana, USA; [b]Texas Transportation Institute, San Antonio, USA; [c]Research Institute of Highway Ministry of Transport, Beijing, China; [d]Department of Civil Engineering, Auburn University, Auburn, Alabama, USA

ABSTRACT

Animal vehicle crash is a critical yet often under-emphasized safety concern of Louisiana. During 2014-2018, over 14,000 animal-related crashes cost Louisiana more than $520 million. To identify multiple key contributing factors and their association patterns, this study applied association rules mining in the dataset of animal-related roadway crashes that occurred during 2014-2018. Since high proportions of animal-related crashes involve complaint and no injury of vehicle occupants, separate analyses were performed for KAB (fatal, severe, and moderate injury) and CO (possible/complaint and no injury) crashes. Top rules ordered by higher lift values were interpreted and compared to implicate the quantified likelihood of crash patterns. KAB rules presented the likelihood of associations of characteristics such as unlighted dark conditions, interstate and parish roads, a wide range of speed limits, residential and open country locations, normal and rainy weather conditions, light trucks, young drivers, etc. The majority of CO crash patterns were associated with interstates, straight segments, normal driver conditions, clear weather, unlighted dark conditions, open country locations, a speed limit of 97 km/h or higher, etc. Findings in this study and their implications supported by prior studies are expected to be beneficial in strategic planning for identifying implementable countermeasures for animal-vehicle crashes.

ARTICLE HISTORY
Received 11 July 2021
Accepted 12 September 2022

KEYWORDS
Animal-vehicle crashes; association rule mining; unsupervised learning; rural roads; roadway lighting; open country location; nighttime crashes

Introduction

Crashes due to animals on the roadway are considerably critical safety and economic concern (Huijser et al., 2007). These crashes are often portrayed as a danger both for public health and wildlife safety (Cherry et al., 2019). Since the majority of animal-vehicle crashes are non-fatal, they are often under-recognized (Ang et al., 2019). However, long-term crash data showed that fatalities due to motor vehicle crashes involving collisions with animals in the United States are generally on the rise (Insurance Institute for Highway Safety, 2020). Apart from hit-animal crashes, roadway departure and collision with other vehicles also occur due to abrupt maneuvers of vehicles trying to avoid animals on the roadway. Research on hit-animal crashes on roadway using nationwide fatal (Khattak, 2003; Langley et al., 2006; Sullivan, 2009, 2011) and non-fatal (Conn et al., 2004) crash data have been conducted to present descriptive statistics and trend of crash characteristics. Studies on location-based specific types of the animal-vehicle crash also exist specifically in states where animal crashes are a primary area of roadway safety focus (e.g. Garrett & Conway, 1999; Savolainen & Ghosh, 2008).

The state of Louisiana has been identified as a moderate risk state in terms of insurance claims due to collisions with animals (1 in 169 claims) during July 2018 - June 2019 (StateFarm, 2019). Apart from crashes with wild animals such as deer, alligators, etc., there have been reports of crashes with farm animals (Farm Sanctuary, 2006; KPLC 7 News, 2014). Due to a lack of previous research on this topic in this state, it is important to conduct research analyzing the characteristics of crashes related to animals on the roadway.

It is well-known that crashes with animals especially wildlife are prevalently single-vehicle crashes and mainly occur on rural two-lane highways (M. P. Huijser et al., 2004). A Louisiana study estimates that the odds of a roadway departure due to animal presence was found to be very high (odds ratio of 17.187) compared to normal roadway conditions on Louisiana rural two-lane highways (Rahman et al., 2021). In addition to colliding with animals, drivers get surprised by unexpectedly animal(s) running across the road specifically in rural areas and lose control of the vehicle causing a roadway departure. Although Louisiana is consistently known as a dangerous state for animal-vehicle crashes among the central south states, no research analyzing Louisiana's hit-animal crash characteristics could be found.

According to the crash cost estimates from the Louisiana Department of Transportation and Development (DOTD) (LaDOTD, 2016), a total of 14,616 crashes involving animals on the roadway in 2014-2018 cost Louisiana more than

CONTACT M. Ashifur Rahman ashifur@louisiana.edu Department of Civil Engineering, University of Louisiana, Lafayette, Louisiana, USA

$520 million. A vast majority (84%) of those crashes were no injury crashes indicating mainly vehicle and other property damages. Because the majority of crashes due to animals on the roadway involve no evidently visible bodily injury of drivers or passengers, two separate analyses of crashes with different severity levels were performed. For analyzing animal-involved crashes in Louisiana, the authors have undertaken a data mining technique known as 'Association Rules Mining'.

Data mining is a novel approach to understanding non-trivial valid patterns in a large dataset by involving statistical learning, modeling theories, algorithm potential, and effective database management (Piatetsky-Shapiro et al., 1996). As conventional statistical parametric models establish relationships based on independent and dependent variable distributions, they exhibit a restrained ability to untangle complex associations among multiple variables in high-dimensional datasets (Mannering & Bhat, 2014). These shortcomings increase with a growing number of variables that lead to invalid results due to the scatterings. Association rule mining – an unsupervised learning method utilized across a set of divergent fields like market basket analysis, product recommendation, medical diagnosis in revealing interesting and non-trivial relations of variables – has currently emerged as a popular approach in highway safety analysis (Das et al., 2018; Das et al., 2020; Pande & Abdel-Aty, 2009; Rahman et al., 2021; Weng et al., 2016). Several transportation safety domains have been explored using association rule mining to explore hidden patterns in the areas of interest. These include crash studies using information related to crash location, crash manners, driver as reported by police agencies; naturalistic driving studies that are expected to provide insight in everyday driver behavior in addition ro driving locations, observational studies such as roadside obsevation data. To name a few on the latest transportation safety studies are prevalence of cellphone distraction by roadside observations (Rahman et al., 2021), roadway departure on rural two-lane highways (Rahman et al., 2021), investigation of fatal and injury crash patterns of teen drivers (Hossain et al., 2021), patterns of near-crash events in a naturalistic driving dataset (Kong et al., 2021), etc. The potential of association rule mining for utilizing its capability in identifying coexisting crash characteristics for decision-making by road safety agencies by targeting specific crash issues (Feng et al., 2020; Pande & Abdel-Aty, 2006). Based on predefined support and confidence threshold, association rules can describe relationships between variables in numerous circumstances without restricting the nature of variables (independent or dependent). Besides, association rule mining can deal with both large and small datasets compared to other non-parametric data mining approaches (Montella et al., 2012).

Considering the advantages of association rule mining over conventional statistical approach in handling a dataset with a large number of variables and the flexibility to interpret variable relationships without any predetermined assumptions and hypotheses, association rule mining has been utilized in this study to identify the crash categories (i.e. crash contributory factors and characteristics) that occur together in animal-related crashes. Based on the understanding of the KABCO injury scale (K - fatal injury, A - severe injury, B - moderate injury, C - possible/complaint injury, O - no injury), the analysis in this study was performed in two different severity levels for two reasons – 1) since high proportions animal-related crash involve complaint and no injury of vehicle occupants (i.e. driver and passenger), these crashes are expected to be characteristically different; and 2) in case of a collision with an animal, KAB crashes would have a higher impact force compared to CO crashes. Specifically, the researchers set the objectives of this study to:

- Identify associative factors of KAB and CO crashes that occurred due to animals on the roadway.
- Compare the pattern of these two types of crashes.

Methodology

Dataset

Animal crash data were initially collected from the online password-protected repository of crashes, 'Crash 1′ maintained by the Louisiana DOTD (LADOTD, 2021). Then some additional crash characteristics data were collected from annually developed more comprehensive crash data files in Microsoft Access format. To reflect recent crashes due to the presence of the animal on roadway, the following criteria in five years (2014-2018) crash data were included:

- Crash type: collision with an animal
- Roadway condition: animal on the roadway

The distribution of animal-related crashes by injury severity is – fatal: 4 (0.02%), severe: 43 (0.29%), moderate: 451 (3.09%), complaint: 1,853 (12.68%), and no injury: 12,265 (83.91%). Figures 1 and 2 illustrate the descending-ordered absolute frequency of the categories (i.e. crash characteristics) in the two datasets. Initial visual comparison of these two figures also shows the absolute frequency of possible/no injury (CO) crashes is clearly higher compared to confirmed injury (KAB) crashes.

The prevalent normal and presumably safe driving conditions of animal-related crashes in both groups are straight segments, clear weather, and normal driver condition. Expectedly, passenger cars are also involved in the majority of crashes due to high ownership of them. However, the dark unlighted condition appears to be among the important factors that might be influential to animal crashes. Review from previous studies suggested estimated animal-vehicle collisions occurring in dark conditions could be up to 90% (Langley et al., 2006). Crashes occur more frequently during the Fall and early Winter seasons due to the high activity levels of animals such as deer (Khattak, 2003; Sullivan, 2011). CO crashes were prevalent during Fall and Spring (Figure 2), whereas a reasonably large number of animal-related crashes also occurred during Spring that

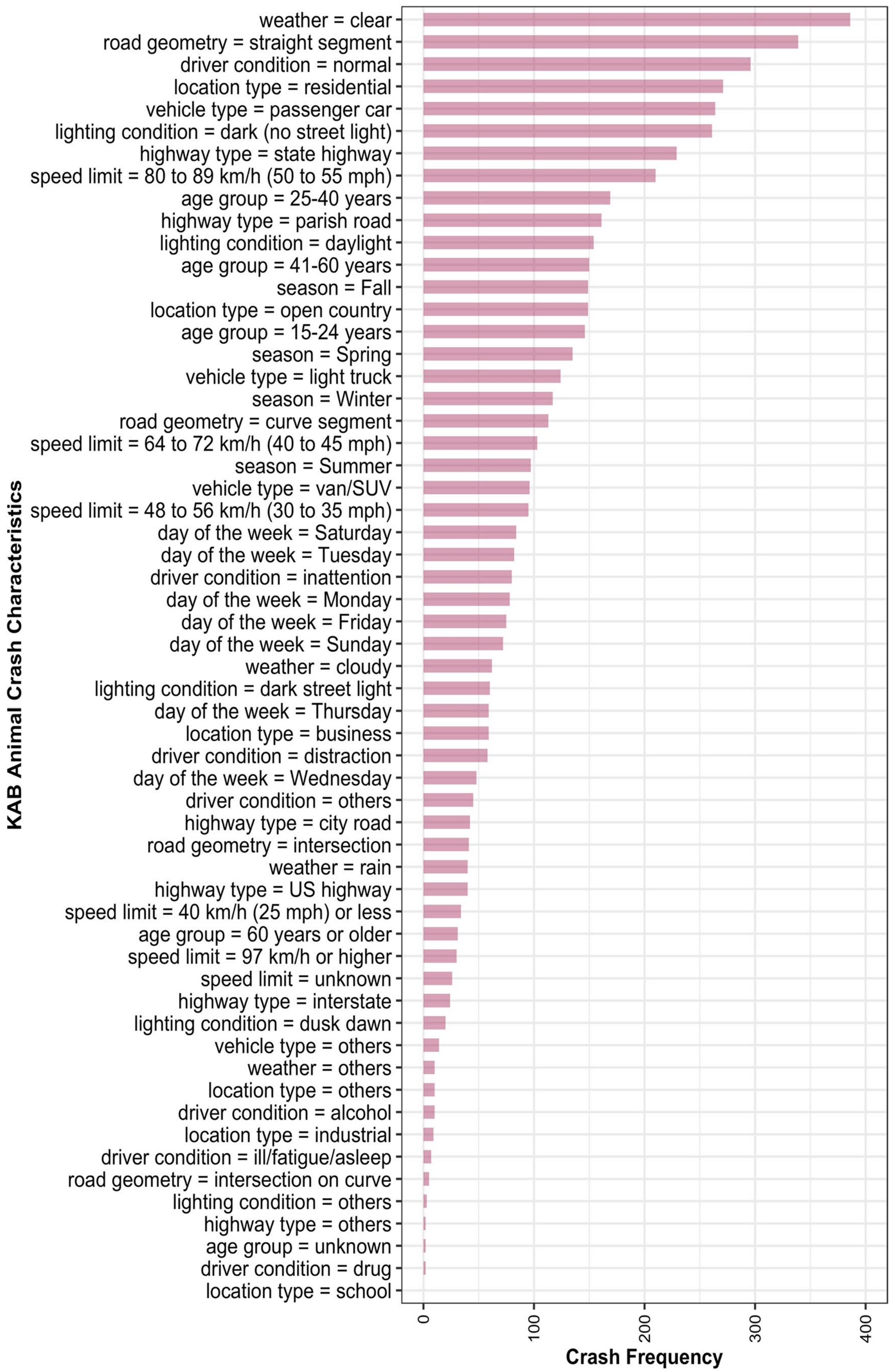

Figure 1. Frequency of KAB crash characteristics.

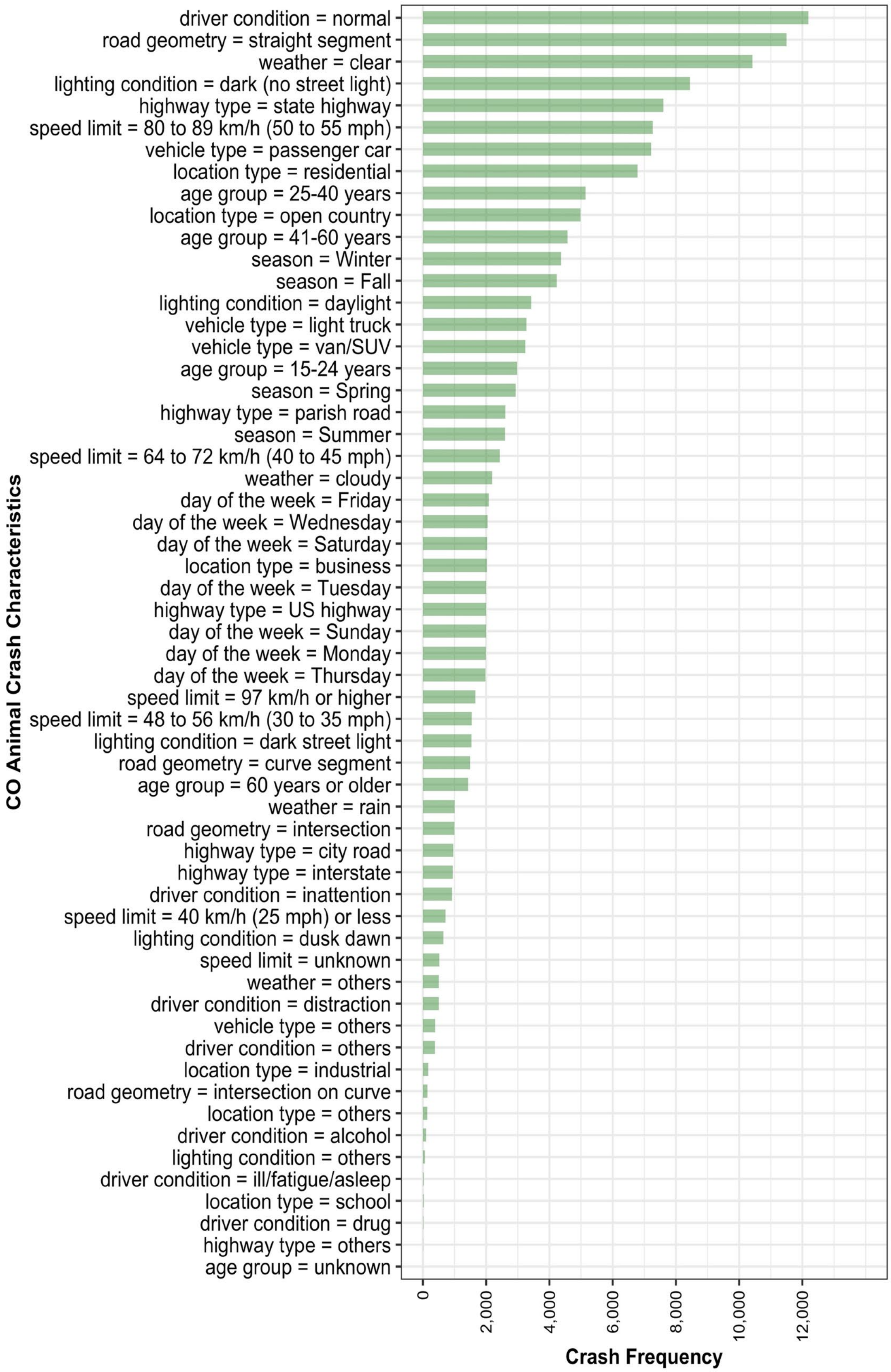

Figure 2. Frequency of CO crash characteristics.

resulted in KAB crashes (Figure 1). Animal-involved crashes occurred more on state highways and highways with a speed limit of 80-89 km/h than other highways. Area type is expected to be immensely influential in animal crashes, as the data shows the majority of crashes in both severity groups occur in a rural setting (either open country or residential area) compared to a business or school area. A thorough look at the crash locations reveals that the residential areas are mainly rural low-density residential areas. To avoid repetitive rules in association rule mining, variables with a very large proportion of one category have not been included in the dataset. For example, animal-related crashes are predominantly single-vehicle crashes; therefore, 'manner of collision' was not included in the analysis. For similar reasons, the variable 'area type' was categorized based on surrounding land use 'open country', 'residential', 'business', 'industrial', 'others', rather than dividing into simple 'urban', and 'rural' categories.

Association rule mining

The basic theoretical background of association rule mining is as follows:

Let $T=\{t_1,t_2,t_3,\ldots,t_n\}$ be a set of individual crash information in the database consisting and each crash record in T contains a subset of items (a set of variable categories) in itemset, $I=\{i_1,i_2,i_3,\ldots,i_n\}$. A rule has the form $X \rightarrow Y$, where $X,Y \subseteq I$ and $X \cap Y = \varnothing$, where, X is the antecedent (left-hand side – LHS) and Y is the consequent (right-hand side – RHS). In a n-itemset rule, multiple items can be considered as antecedent and consequent. For example, a 3-itemset rule consists of either two antecedents and one consequent or two consequents and one antecedent. The rules were filtered by three parameters – support, confidence, and lift. Support of any rule can be determined by how frequently that rule or pattern ($X \rightarrow Y$) occurs in the entire dataset. Confidence is the proportion of how recurring $X \rightarrow Y$ together by the number of times X occurs in the dataset. The third parameter 'lift' helps to determine statistical dependence of rule $X \rightarrow Y$ by specifying how more often items are part of the same independent crash events (Brin et al., 1997). The equations of parameters are listed below:

$$Support(X)=\frac{\#X}{N} \tag{1}$$

$$Support(Y)=\frac{\#Y}{N} \tag{2}$$

$$Support(X \rightarrow Y)=\frac{\#(X \cap Y)}{N} \tag{3}$$

$$Confidence(X \rightarrow Y)=\frac{Support(X \rightarrow Y)}{Support(X)} \tag{4}$$

$$Lift(X \rightarrow Y)=\frac{Support(X \rightarrow Y)}{Support(X) \times Support(Y)} \tag{5}$$

where, N = number of crashes, $\#X$ = frequency of occurrences with the antecedent (X), $\#Y$ = frequency of occurrences with consequent (Y), $\#(X \cap Y)$ = frequency of occurrences with both the antecedent (X) and the consequent (Y).

In association rule discovery, the lift value is critical to determine the strength of any rule rather than support and confidence since it represents co-occurrence of antecedent on the conditional likelihood of consequent (Besharati & Tavakoli Kashani, 2018; Pande & Abdel-Aty, 2009). The lift value of any rule higher than 1 specifies positive interrelation between the antecedent and the consequent whereas a value less than 1 indicates a negative interrelation. If the lift value is close to 1, the antecedent will be independent in the probability of the consequent.

Studies developed a variety of algorithms to identify frequent itemset patterns more efficiently based on the specific dataset format (Das et al., 2018; Maimon & Rokach, 2010). Among them, the 'apriori' algorithm was developed for searching reoccurring itemsets with predefined threshold values by presuming the subset of itemsets as frequent (Agrawal et al., 1993). Using a 'bottom-up' method in the apriori algorithm, frequent subsets are expanded one item at a time to conduct candidate generation, and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Apriori uses breadth-first search to identify candidate itemsets efficiently. Details of the algorithm can be found in Agarwal and Srikant (1994).

Results and discussions

For generating the optimum number of interesting and meaningful rules, minimum support of 0.05 and minimum confidence of 0.7 were specified through an iterative approach in line with the previous studies (Das et al., 2018; Pande & Abdel-Aty, 2009; Rahman et al., 2021). Using injury-specific subsets of data, the numbers of rules generated are the following–

- 2-itemset rules: KAB – 40 rules, CO – 59 rules
- 3-itemset rules: KAB – 253 rules, CO – 387 rules
- 4-itemset rules: KAB – 319 rules, CO – 625 rules

To obtain strongly associated co-occurring characteristics in animal crashes, the top ten rules with high lift values in each itemset were extracted. The top ten association rules ordered for 2-, 3-, and 4-itemsets are presented in Tables 1, 2, and 3, respectively. It is important to mention that these rules imply only associations between sets of items rather than direct causation.

2-Itemset rules

The top ten 2-itemset rules ordered from higher to lower lift values are presented in Table 1. Some key takeaways from the interesting rules in terms of lift values are:

- Rule CO#1 indicates that the likelihood of a CO crash in presence of animals on the interstates with a speed limit of 97 km/h or higher is likely to happen 7.809 times in all CO crashes in presence of animals on the interstates. Similarly, rule KAB#1 indicates on state highways, KAB crashes due to the presence of animals are 1.771 times if the speed limit is between 80 to 88 km/h. One study on fatal animal-vehicle crashes estimated that for every mile-per-hour increment in speed over the speed limit there was an increase in the odds of a crash in the darkness of about 2.5% (Sullivan, 2009).

- In unlighted dark conditions, a CO crash is 1.205 times more likely when it occurs on highways with 97 km/h (60 mph) or higher speed limit (Rule CO#7) and 1.221 times more likely in an open countryside location (Rule CO#6). According to some earlier studies, the high likelihood could be attributed to the population of wildlife near the highway and speeding (M. P. Huijser et al., 2004), animal movements during nighttime (Khattak, 2003).
- Even if the driver condition is normal (no influence of alcohol, distraction, fatigue, etc.), KAB crashes due to animal presence would be 1.304 times more likely during rainy weather (rule KAB #7). Due to wet pavement surface, hit animal crashes are more likely due to small braking distance, and swerve maneuvers to avoid animal(s) on the roadway during rainy weather could also lead to severe injury crashes.
- Among the roadway segments which are continuously straight (i.e. no cross-sectional changes such as intersection, curve, or both), KAB crashes with animal(s) are 1.42 times more likely if the speed limit is 97 km/h (60 mph) or higher (rule KAB #3). In the case of CO crashes, the likelihood is lower, 1.146 times (rule CO#8). According to earlier research, animal-vehicle crashes are more likely to happen on higher speed limits (Gunther et al., 1998). As per common perception, crashes due to animal presence should be less likely on straight segments due to better visibility, however, drivers tend to speed on a straight segment of roadways (Barrientos & Bolonio, 2009). Additionally, cross-sectional changes on roadway configuration such as curves are expected to draw extra attention to drivers (Charlton, 2007; Elvik, 2019) to react to animal movement.

3-Itemset rules

The 3-itemset association rules are presented in Table 2. Some of the key observations from the table are:

- Six out of the top ten rules regarding KAB crashes have *highway type = parish road* as consequent. The antecedents always include *speed limit = 48 to 56 km/h (30 to 35 mph)* in all of these rules with *highway type = parish road* as consequent, as the majority of the parish roads have a speed limit of 48 to 56 km/h (30 to 35 mph). For a similar reason, *highway type = state highway* and *speed limit = 80 to 88 km/h (50 to 55 mph)* appear to be closely related. A lift value close to 2 or greater in these rules indicates all strong interdependencies.
- All of the top ten rules regarding CO crashes include these three most prevalent concurrent conditions – *highway type = interstate*, *location type = open country*, *speed limit = 97 km/h (60 mph) or higher*. Due to the sudden presence of animals on high-speed roadways, drivers usually fail to avoid collisions. Additional road signs on animal crossing may help in reducing these crashes.
- Drivers aged 15-24 years are likely to be associated with KAB crashes. Young drivers' higher propensity to animal-related crashes compared to other age groups can also be seen in an FHWA national study (M. P. Huijser et al., 2004).
- Comparing rules KAB#17 and CO#20, in the same scenario, *highway type = state highway, location type = open country → speed limit = 80 to 88 km/h (50 to 55 mph)*, KAB crashes (L= 2.063) are more likely than CO crashes (L = 1.688).

4-Itemset rules

The 4-itemset association rules are presented in Table 3. Key findings from this table are:

Table 1. 2-itemset rules in KAB and CO animal crashes.

#	Rule (Antecedent → Consequent) KAB	Support, S (%)	Confidence, C (%)	Lift, L	Rule (Antecedent → Consequent) CO	Support, S (%)	Confidence, C (%)	Lift, L
1	speed limit = 80 to 88 km/h → highway type = state highway	34	81	1.771	highway type = interstate → speed limit = 97 km/h or higher	6	91	7.809
2	speed limit = 40 km/h or lower → location type = residential	6	88	1.621	highway type = interstate → location type = open country	5	83	2.338
3	speed limit = 97 km/h or higher → road geometry = straight segment	6	97	1.42	speed limit = 97 km/h or higher → location type = open country	8	71	2.019
4	highway type = U.S. highway → driver condition = normal	7	83	1.388	speed limit = 80 to 88 km/h → highway type = state highway	41	80	1.482
5	speed limit = 48 to 56 km/h → location type = residential	14	73	1.335	highway type = parish road → location type = residential	13	70	1.461
6	highway type = parish road → location type = residential	23	72	1.324	location type = open country → lighting condition = dark (no streetlight)	26	73	1.221
7	weather = rain → driver condition = normal	6	78	1.304	speed limit = 97 km/h or higher → lighting condition = dark (no streetlight)	8	72	1.205
8	speed limit = 97 km/h or higher → weather = clear	5	90	1.161	speed limit = 97 km/h or higher → road geometry = straight segment	11	93	1.146
9	season = Fall → road geometry = straight segment	22	75	1.104	highway type = interstate → road geometry = straight segment	6	91	1.116
10	highway type = U.S. highway → road geometry = straight segment	6	75	1.102	speed limit = 97 km/h or higher → driver condition = normal	11	94	1.089

Table 2. 3-itemset rules in KAB and CO animal crashes.

#	Rule (Antecedent → Consequent) KAB	Support, S (%)	Confidence, C (%)	Lift, L	Rule (Antecedent → Consequent) CO	Support, S (%)	Confidence, C (%)	Lift, L
11	speed limit = 48 to 56 km/h, vehicle type = light truck → highway type = parish road	5	93	2.864	highway type = interstate, location type = open country → speed limit = 97 km/h or higher	5	96	8.228
12	speed limit = 48 to 56 km/h, lighting condition = dark (no streetlight) → highway type = parish road	8	88	2.733	highway type = interstate, road geometry = straight segment → speed limit = 97 km/h or higher	6	94	8.036
13	location type = residential, speed limit = 48 to 56 km/h → highway type = parish road	11	80	2.466	highway type = interstate, driver condition = normal → speed limit = 97 km/h or higher	6	94	8.005
14	speed limit = 48 to 56 km/h, age group = 15-24 y → highway type = parish road	5	76	2.365	highway type = interstate, road geometry = straight segment → location type = open country	5	86	2.441
15	road geometry = straight segment, speed limit = 48 to 56 km/h → highway type = parish road	9	73	2.268	highway type = interstate, driver condition = normal → location type = open country	5	86	2.424
16	speed limit = 48 to 56 km/h, driver condition = normal → highway type = parish road	7	71	2.191	speed limit = 97 km/h or higher, lighting condition = dark (no streetlight) → location type = open country	7	77	2.188
17	highway type = state highway, location type = open country → speed limit = 80 to 88 km/h	13	87	2.063	road geometry = straight segment, speed limit = 97 km/h or higher → location type = open country	8	72	2.043
18	season = Summer, speed limit = 80 to 88 km/h → highway type = state highway	7	93	2.012	weather = clear, speed limit = 97 km/h or higher → location type = open country	6	72	2.043
19	highway type = state highway, lighting condition = dark (no streetlight) → speed limit = 80 to 88 km/h	21	85	2.011	speed limit = 97 km/h or higher, driver condition = normal → location type = open country	8	72	2.034
20	season = Spring, highway type = state highway → speed limit = 80 to 88 km/h	10	84	1.997	highway type = state highway, location type = open country → speed limit = 80 to 88 km/h	17	87	1.688

- The factor *lighting condition = dark (no streetlight)* appears to be influential in confirmed injury crashes both on parish road (KAB#21, KAB#22, KAB#23) and on state highway (KAB#28, KAB#30). The combined effect of *day of the week = Saturday* (weekend) and *lighting condition = dark (no streetlight)* can be seen in rule KAB#30. Improvement of lighting conditions can reduce these crashes.
- A particular scenario, *highway type = state highway, location type = open country, lighting condition = dark (no streetlight) → speed limit = 80 to 88 km/h (50 to 55 mph)*, is more associated with KAB crashes (KAB#28) than CO crashes (CO#30).
- In 4-itemset rules, interestingly, on parish roads with a speed limit of 48-56 km/h (30 to 35 mph), KAB crashes were found to be associated with not only unlighted dark conditions (KAB#21, KAB#22, KAB#23), but also other perceived normal condition such as straight segment (KAB#24), clear weather (KAB#25), normal condition of the driver (KAB#26) especially in the rural residential areas (KAB#24, KAB#25, KAB#26). This could still be linked with drivers' tendency to speed up in low-speed rural areas, indicating a need for speed enforcement.

Discussion by crash characteristics

Time and condition related characteristics

Darkness was found to play a significant role in both KAB and CO animal-related crashes, similar to the findings of previous studies (M. P. Huijser et al., 2004; Khattak, 2003; Langley et al., 2006; Sullivan, 2009). This implicates that compromised visibility due to the absence of streetlights in reacting to sudden movement of animals in darkness remains a critical issue for drivers regardless of injury severity types. Although the Fall season has been more frequently linked to both KAB and CO crashes (Figures 1 and 2), all three seasons could be linked to only KAB crashes in top rules. This could vary by location due to animal movements and activities impacted by temperature intensity and trend of animal migration (Al-Bdairi et al., 2020; Hedlund et al., 2004). Both KAB and CO crashes peaked on Saturday (Figures 1 and 2); however, the association of Saturday with darkness, state highways, and a speed limit of 80 to 88 km/h can only be seen in one KAB rule among all 30 rules presented. This finding is partially supported by an Italian study that showed significantly high crashes during the weekend (Putzu et al., 2014).

Roadway and location characteristics

Both KAB and CO crashes with animal presence occur mostly in open country areas. With the association rule mining, it was also identified that KAB crashes associated with animal presence could occur on parish roads, state highways in addition to interstates. Coinciding with those results, KAB crashes were found to be associated with roadways with a range of speed limits – 48 to 56 km/h, 80-88 km/h, and 97 km/h or higher, respectively. CO crashes were mainly associated with roadways with 80-88 km/h and 97 km/h speed limits. It has been implicated that the risk

Table 3. 4-itemset rules in KAB and CO animal crashes.

#	Rule (Antecedent → Consequent) KAB	Support, S (%)	Confidence, C (%)	Lift, L	Rule (Antecedent → Consequent) CO	Support, S (%)	Confidence, C (%)	Lift, L
21	location type = residential, speed limit = 48 to 56 km/h, lighting condition = dark (no streetlight) → highway type = parish road	7	92	2.842	highway type = interstate, location type = open country, road geometry = straight segment → speed limit = 97 km/h or higher	5	97	8.292
22	weather = clear, speed limit = 48 to 56 km/h, lighting condition = dark (no streetlight) → highway type = parish road	6	90	2.794	highway type = interstate, location type = open country, driver condition = normal → speed limit = 97 km/h or higher	5	97	8.278
23	road geometry = straight segment, speed limit = 48 to 56 km/h, lighting condition = dark (no streetlight) → highway type = parish road	5	89	2.762	highway type = interstate, road geometry = straight segment, driver condition = normal → speed limit = 97 km/h or higher	5	95	8.135
24	location type = residential, road geometry = straight segment, speed limit = 48 to 56 km/h → highway type = parish road	7	79	2.435	weather = clear, speed limit = 97 km/h or higher, lighting condition = dark (no streetlight) → location type = open country	5	78	2.206
25	weather = clear, location type = residential, speed limit = 48 to 56 km/h → highway type = parish road	8	78	2.413	road geometry = straight segment, speed limit = 97 km/h or higher, lighting condition = dark (no streetlight) → location type = open country	6	78	2.197
26	location type = residential, speed limit = 48 to 56 km/h, driver condition = normal → highway type = parish road	6	74	2.279	speed limit = 97 km/h or higher, driver condition = normal, lighting condition = dark (no streetlight) → location type = open country	6	77	2.177
27	highway type = state highway, location type = open country, vehicle type = passenger car → speed limit = 80 to 88 km/h	7	95	2.243	weather = clear, road geometry = straight segment, speed limit = 97 km/h or higher → location type = open country	6	73	2.076
28	highway type = state highway, location type = open country, lighting condition = dark (no streetlight) → speed limit = 80 to 88 km/h	9	94	2.226	weather = clear, speed limit = 97 km/h or higher, driver condition = normal → location type = open country	6	73	2.069
29	weather = clear, road geometry = straight segment, speed limit = 48 to 56 km/h → highway type = parish road	7	72	2.219	road geometry = straight segment, speed limit = 97 km/h or higher, driver condition = normal → location type = open country	7	73	2.055
30	day of the week = Saturday, highway type = state highway, lighting condition = dark (no streetlight) → speed limit = 80 to 88 km/h	5	93	2.196	highway type = state highway, location type = open country, lighting condition = dark (no streetlight) → speed limit = 80 to 88 km/h	13	88	1.706

of animal crashes in darkness could vary posted speed limit (Sullivan, 2011). This study suggests, besides darkness (KAB#12, KAB#23), the likelihood of KAB crashes on a relatively low speed limit range (48-56 km/h) that typically occur on parish roads could vary based on their association with the residential area (KAB#5), light truck (KAB#11) and young driver (15-24 years) (KAB#14). However, on a very high-speed roadway with a speed limit of 97 km/h (60 mph) or higher during normal conditions, KAB crashes are more likely than CO crashes. Some studies suggested high concentration of animal crashes could be reduced following a lowering of the speed limit of highways (Danks & Porter, 2010; Gunson et al., 2011). Results from the rules suggest low-density residential areas have been associated with KAB crashes in darkness, whereas open country areas with high speed limits have been found to be associated with CO crashes where darkness was not a factor. Several of these rules of KAB crashes with high lift values include darkness indicating parish and roads in low-density residential areas may encourage more animal activities such as crossing streets which could end up in KAB crashes. In contrast, high-speed roadways (i.e. interstates) may encourage less animal activity by noisy and more frequent traffic. A partially similar result has been found by Al-Bdairi et al. (2020) who interpreted that rural freeways may not be a significant factor for severer injury in animal-vehicle crashes in recent years.

Driver and vehicle characteristics

No driver and vehicle characteristics have been featured in the top 10 rules of 2-, 3-, and 4-itemset rules of CO crashes. Only two KAB rules could be found that included light trucks (KAB#11) and young drivers aged 15-24 years (KAB#14). Although middle-aged drivers are proportionately more involved in animal crash frequencies compared to young drivers (Figure 1), the distribution of such crashes by vehicle miles traveled could still be higher for young drivers (M. P. Huijser et al., 2004). Rules identified in this study showed that animal crashes with young drivers at fault were associated with a speed limit between 48 to 56 km/h on parish roads, which could very well be attributed to their speeding tendency, overestimation of driving abilities, less skill in performing emergency maneuvers, etc. (FHWA, 2001). Light truck or pick-ups are also associated with KAB crashes on parish roads with a speed limit of 48-56 km/h, indicating a higher likelihood of these crashes in rural areas. KAB crashes with light trucks occur more in rural areas (NHTSA, 2021), where ownership of these vehicles is higher (Anderson et al., 2001; Cargurus, 2021).

Conclusion

Since animal-related crashes resulting in either no or possible injury are disproportionately high (96.6%) in Louisiana, there is less emphasis on animal-vehicle crashes. There is a clear research gap in understanding the collective associations of characteristics in crashes due to animal presence on roadways. Non-parametric data mining methods can successfully identify the interactions between different features by overcoming the predetermined assumptions of statistical modeling. This study used an unsupervised learning method, association rule mining, to mitigate the research gap. It is perhaps the first broad analysis of animal-related crashes in Louisiana identifying the crash characteristics and discreetly stipulating the need for further attention.

The study clearly distinguishes the animal crashes that end up in KAB severity or CO severity by comparing the top rules segregated by 2-itemsets, 3-itemsets, and 4-itemsets allowing for multiple concurrent crash characteristics in the results. Rules of similar associated characteristics between these two severity groups allow researchers to compare the quantified measures such as lift values. Top rules ordered by higher lift values were interpreted and compared to implicate the likelihood of animal crash patterns presenting collective crash characteristics. Although several rules associate KAB crashes across various speed limit groups, the likelihood of these crashes in high speed limit (97 km/h or higher) is higher compared to CO crashes. KAB rules also presented the likelihood of collective combinations of characteristics such as darkness, interstate and parish roads, a wide range of speed limits, residential and open country locations, normal and rainy weather conditions, light truck, young driver, etc. The majority of the CO crash patterns were associated with interstates, straight segments, normal conditions of drivers, clear weather, unlighted dark conditions, open country locations, a speed limit of 97 km/h or higher, etc. Implications of the findings supported by available prior studies were presented by time and condition characteristics, roadway and location characteristics, and driver and vehicle characteristics. These implications are expected to benefit transportation researchers and planners in planning for implementable countermeasures for animal-vehicle crashes.

Considering the high frequency of no injury crashes, strategic implementation of economically feasible countermeasures for crashes due to animal presence requires a thorough understanding of co-occurring crash characteristics. The factors identified in this study indicate that countermeasures such as speed management, retroreflective warning signs, and animal crossing signs can be considered useful in the absence of streetlights on rural roadways. Driver education encompassing avoiding collisions with animals and minimizing injury in animal-vehicle crashes, by understanding the animal warning signs, etc. especially to young drivers could also be beneficial. Continuous data-driven analysis of animal-vehicle crashes is also necessary for the implementation and evaluation of these crash countermeasures.

Acknowledgment

The authors would like to thank the Louisiana DOTD for providing the data. This research did not receive any grant from any funding agency. The authors appreciate the comments from the two anonymous reviewers. The authors would also like to thank Mary Kathryn Sevin for her assistance in preparing the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

M. Ashifur Rahman http://orcid.org/0000-0001-6940-1599
Subasish Das http://orcid.org/0000-0002-1671-2753
Xiaoduan Sun http://orcid.org/0000-0001-7282-1340
Ming Sun http://orcid.org/0000-0002-1032-4135
Md. Mahmud Hossain http://orcid.org/0000-0002-2737-6951

References

Agarwal, R., & Srikant, R. (1994). *Fast algorithms for mining association rules* [Paper presentation]. 20th International Conference on Very Large Data Bases, VLDB (Vol. 1215, pp. 487–499). https://doi.org/10.1345/aph.1K157

Agrawal, R., Imieliński, T., & Swami, A. (1993). *Mining association rules between sets of items in large databases 1993 ACM SIGMOD international conference on Management of data* (pp. 207–216). https://doi.org/10.1145/170035.170072.

Al-Bdairi, N. S. S., Behnood, A., & Hernandez, S. (2020). Temporal stability of driver injury severities in animal-vehicle collisions: A random parameters with heterogeneity in means (and variances) approach. *Analytic Methods in Accident Research*, *26*, 100120. https://doi.org/10.1016/j.amar.2020.100120

Anderson, C. L., Agran, P. F., & Winn, D. G. (2001). Pickup truck use in the National Personal Transportation Survey. *Accident; Analysis and Prevention*, *33*(4), 499–506. https://doi.org/10.1016/S0001-4575(00)00063-4

Ang, J. Y., Gabbe, B., Cameron, P., & Beck, B. (2019). Animal–vehicle collisions in Victoria, Australia: An under-recognised cause of road traffic crashes. *Emergency Medicine Australasia : EMA*, *31*(5), 851–855. https://doi.org/10.1111/1742-6723.13361

Barrientos, R., & Bolonio, L. (2009). The presence of rabbits adjacent to roads increases polecat road mortality. *Biodiversity and Conservation*, *18*(2), 405–418. https://doi.org/10.1007/s10531-008-9499-9

Besharati, M. M., & Tavakoli Kashani, A. (2018). Which set of factors contribute to increase the likelihood of pedestrian fatality in road crashes? *International Journal of Injury Control and Safety Promotion*, *25*(3), 247–256. https://doi.org/10.1080/17457300.2017.1363781

Brin, S., Motwani, R., & Silverstein, C. (1997). Beyond market baskets: Generalizing association rules to dependence rules. *ACM SIGMOD Record*, *26*(2), 265–276. https://doi.org/10.1023/A:1009713703947

Cargurus. (2021). *2020 pickup truck sentiment study: Measuring COVID-19's impact on the category*. Retrieved from https://www.cargurus.com/Cars/aboutCargurus.html

Charlton, S. G. (2007). The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments. *Accident; Analysis and Prevention*, *39*(5), 873–885. https://doi.org/10.1016/J.AAP.2006.12.007

Cherry, C. C., Dietz, S., Sauber-Schatz, E., Russell, S., Proctor, J., & Buttke, D. (2019). Characteristics of animal-related motor vehicle crashes in select National Park Service units—United States, 1990–2013. *Traffic Injury Prevention*, *20*(1), 58–63. https://doi.org/10.1080/15389588.2018.1508835

Conn, J. M., Annest, J. L., & Dellinger, A. (2004). Nonfatal motor-vehicle animal crash-related injuries – United States, 2001-2002. *Journal of Safety Research*, *35*(5), 571–574. https://doi.org/10.1016/j.jsr.2004.10.002

Danks, Z. D., & Porter, W. F. (2010). Temporal, spatial, and landscape habitat characteristics of moose-vehicle collisions in western Maine. *The Journal of Wildlife Management*, *74*(6), 1229–1241. https://doi.org/10.1111/j.1937-2817.2010.tb01243.x

Das, S., Dutta, A., Jalayer, M., Bibeka, A., & Wu, L. (2018). Factors influencing the patterns of wrong-way driving crashes on freeway exit ramps and median crossovers: Exploration using 'Eclat' association rules to promote safety. *International Journal of Transportation Science and Technology*, *7*(2), 114–123. https://doi.org/10.1016/j.ijtst.2018.02.001

Das, S., Sun, X., Goel, S., Sun, M., Rahman, M. A., & Dutta, A. (2020). Flooding related traffic crashes: Findings from association rules. *Journal of Transportation Safety and Security*, *0*(0), 1–19. https://doi.org/10.1080/19439962.2020.1734130

Elvik, R. (2019). The more (sharp) curves, the lower the risk. *Accident; Analysis and Prevention*, *133*, 105322. https://doi.org/10.1016/J.AAP.2019.105322

Farm Sanctuary. (2006). *Highway accidents involving farm animals*. Retrieved from https://thehill.com/sites/default/files/HighwayAccidentsFarmAnimals_0.pdf

Feng, M., Zheng, J., Ren, J., & Xi, Y. (2020). Association rule mining for road traffic accident analysis: A case study from UK. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11691 LNAI* (pp. 520–529). https://doi.org/10.1007/978-3-030-39431-8_50

FHWA. (2001). *Young drivers and highway design and operations: Findings and recommended research*. Federal Highway Administration, Washington, DC. Publication No. FHWA-RD-01-120, Retrieved from https://rosap.ntl.bts.gov/view/dot/40962/dot_40962_DS1.pdf

Garrett, L. C., & Conway, G. A. (1999). Characteristics of moose-vehicle collisions in anchorage, Alaska, 1991–1995. *Journal of Safety Research*, *30*(4), 219–223. https://doi.org/10.1016/S0022-4375(99)00017-1

Gunson, K. E., Mountrakis, G., & Quackenbush, L. J. (2011, April 1). Spatial wildlife-vehicle collision models: A review of current work and its application to transportation mitigation projects. *Journal of Environmental Management*, *92*(4), 1074–1082.https://doi.org/10.1016/j.jenvman.2010.11.027

Gunther, K. A., Biel, M. J., & Robison, H. L. (1998). *Factors influencing the frequency of road-killed wildlife in Yellowstone National Park*. Proceedings of International Conference on Wildlife Ecology and Transportation (ICOWET 1998). Fort Meyers, Florida.

Hedlund, J. H., Curtis, P. D., Curtis, G., & Williams, A. F. (2004). Methods to reduce traffic crashes involving deer: What works and what does not. *Traffic Injury Prevention*, *5*(2), 122–131. https://doi.org/10.1080/15389580490435079

Hossain, M. M., Sun, X., Mitran, E., & Rahman, M. A. (2021). Investigating fatal and injury crash patterns of teen drivers with unsupervised learning algorithms. *IATSS Research*, *45*(4), 561–573. https://doi.org/10.1016/j.iatssr.2021.07.002

Huijser, M. P., McGowen, P., Fuller, J., Hardy, A., Kociolek, A., Clevenger, A. P., … Ament, R. (2004). *Wildlife-vehicle collision reduction study: Report to congress*. Federal Highway Administration, Washington, DC. Publication No. FHWA-HRT-08-034, Retrieved from https://www.fhwa.dot.gov/publications/research/safety/08034/08034.pdf

Huijser, M. P., Fuller, J., Wagner, M. E., Amanda, H., & Clevenger, A. P. (2007). Animal-vehicle collision data collection. In *NCHRP SYNTHESIS 370*. https://doi.org/10.17226/23138

Insurance Institute for Highway Safety. (2020). Collisions with fixed objects and animals. Retrieved September 23, 2020, from Fatality Facts 2018 website: https://www.iihs.org/topics/fatality-statistics/detail/collisions-with-fixed-objects-and-animals

Khattak, A. J. (2003). Human fatalities in animal-related highway crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *1840*(1), 158–166. https://doi.org/10.3141/1840-18

Kong, X., Das, S., "Tracy" Zhou, H., & Zhang, Y. (2021). Patterns of near-crash events in a naturalistic driving dataset: Applying rules mining. *Accident; Analysis and Prevention*, *161*, 106346. https://doi.org/10.1016/J.AAP.2021.106346

KPLC 7 News. (2014). Fatal accident involving cow on road sheds light on livestock owners' responsibilities. In *KPLC virtual and VHF digital channel 7*. United States

LaDOTD. (2016). Louisiana-specific cost of crashes. *Highway Safety Analysis Toolbox*. Retrieved July 12, 2020, from http://wwwsp.dotd.la.gov/Inside_LaDOTD/Divisions/Multimodal/Highway_Safety/Pages/Highway_Safety_Analysis_Toolbox.aspx

LADOTD. (2021). Louisiana Department of Transportation and Development Highway Crash List. Retrieved July 5, 2018, from http://www8.dotd.la.gov/crash1/

Langley, R. L., Higgins, S. A., & Herrin, K. B. (2006). Risk factors associated with fatal animal-vehicle collisions in the United States, 1995-2004. *Wilderness & Environmental Medicine*, *17*(4), 229–239. https://doi.org/10.1580/06-WEME-OR-001R1.1

Maimon, O., & Rokach, L. (2010). *Data mining and knowledge discovery handbook*. New York: Springer Science+Business Media, Inc. https://doi.org/10.1007/978-0-387-09823-4

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research*, *1*, 1–22. https://doi.org/10.1016/j.amar.2013.09.001

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

NHTSA (2021). Fatality facts 2019: Urban/rural comparison. *Highway Safety Topics*. Retrieved January 8, 2022, from https://www.iihs.org/topics/fatality-statistics/detail/urban-rural-comparison

Pande, A., & Abdel-Aty, M. (2006). Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways. *Transportation Research Record: Journal of the Transportation Research Board*, *1953*(1), 31–40. https://doi.org/10.3141/1953-04

Pande, A., & Abdel-Aty, M. (2009). Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. *Safety Science*, *47*(1), 145–154. https://doi.org/10.1016/j.ssci.2007.12.001

Piatetsky-Shapiro, G., Fayyad, U., & Smith, P. (1996). From data mining to knowledge discovery: An overview. *Advances in Knowledge Discovery and Data Mining*, *1*, 35.

Putzu, N., Bonetto, D., Civallero, V., Fenoglio, S., Meneguz, P. G., Preacco, N., & Tizzani, P. (2014). Temporal patterns of ungulate-vehicle collisions in a subalpine Italian region. *Italian Journal of Zoology*, *81*(3), 463–470. https://doi.org/10.1080/11250003.2014.945974

Rahman, M. A., Sun, X., Das, S., & Khanal, S. (2021). Exploring the influential factors of roadway departure crashes on rural two-lane highways with logit model and association rules mining. *International Journal of Transportation Science and Technology*, *10*(2), 167–183. https://doi.org/10.1016/j.ijtst.2020.12.003

Rahman, M. A., Sun, X., Sun, M., & Shan, D. (2021). Investigating characteristics of cellphone distraction with significance tests and association rule mining. *IATSS Research*, *45*(2), 198–209. https://doi.org/10.1016/j.iatssr.2020.09.001

Savolainen, P., & Ghosh, I. (2008). Examination of factors affecting driver injury severity in Michigan's single-vehicle–deer crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *2078*(1), 17–25. https://doi.org/10.3141/2078-03

StateFarm. (2019). Herd mentality: Avoiding animal collisions. Retrieved from https://newsroom.statefarm.com/asset/776174/animal-collision-map-527554

Sullivan, J. M. (2009). *Relationships between lighting and animal-vehicle collisions*. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/64281

Sullivan, J. M. (2011). Trends and characteristics of animal-vehicle collisions in the United States. *Journal of Safety Research*, *42*(1), 9–16. https://doi.org/10.1016/j.jsr.2010.11.002

Weng, J., Zhu, J. Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident; Analysis and Prevention*, *92*, 43–52. https://doi.org/10.1016/j.aap.2016.03.017