Check for updates

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–17
© National Academy of Sciences: Transportation Research Board 2023
Article reuse guidelines: sagepub.com/journals-permissions
DOI: 10.1177/03611981231166688
journals.sagepub.com/home/trr
SAGE

Applying Data Mining Methods to Explore Animal-Vehicle Crashes

M. Ashifur Rahman[1], Subasish Das[2], Julius Codjoe[1], Elisabeta Mitran[1], Xiaoduan Sun[1], Kwabena Abedi[1], and Md Mahmud Hossain[3]

Abstract

Animal-vehicle crashes (AVCs) are a significant issue in Louisiana that requires attention. Data on AVCs that occurred from 2015 to 2020 were obtained from the Louisiana Department of Transportation and Development (DOTD), including 14,349 crashes with major injury (KA), minor injury (BC), and no injury (O) severity groups. Aiming to find the collective association of attributes from AVC data, which are categorical in nature, this study utilized two data mining methods: multiple correspondence analysis (MCA) and association rule mining (ARM). Five hierarchical clusters that were generated from the BC and O AVC datasets were particularly significant. Among several other findings, MCA revealed that BC and O AVCs are more concentrated on parish roads during the spring season, while O AVCs in the fall and winter tend to occur on highways with speed limits of 50 mph or higher. ARM revealed that moderate-speed parish roads are frequently associated with KA and BC AVCs, particularly in residential areas and during the spring season, and they often involve young drivers. The findings of this study can be particularly beneficial by considering the spatiotemporal factors associated with animal concentration and movement to develop targeted interventions and mitigation strategies.

Keywords

animal-vehicle crashes, multiple correspondence analysis, hierarchical clustering, association rule mining

Animal-vehicle crashes (AVCs) remain an important transportation safety issue in the U.S., as both long-term and recent crash records indicate that the frequency of AVCs is increasing (*1*, *2*). Understanding the potential factors of AVCs is not only important from the transportation safety perspective, but also in relation to ecological resilience. It has been suggested that the low fatality rate associated with AVCs may be why they do not receive the same level of attention as other types of crashes (*3*). However, despite largely being nonfatal, the toll of AVCs can be intense for crash victims as the repairs of property damages caused by AVCs are often not compensated for without comprehensive insurance coverage (*4*).

Louisiana has a significantly high incidence of AVCs, as evidenced by a 6% increase in AVCs from 2017 to 2019. The majority of these crashes occur with deer, alligators, and bears (*5*). However, crashes with farm animals have also been reported (*6*). AVCs tend to occur in the fall and winter months, which happen to overlap with the breeding season of deer (*7*). The State Farm Insurance Company, which possesses 32% of Louisiana's market share of auto insurance companies, estimates that the 2022–2023 likelihood of a claim involving animals in Louisiana is 1 in 178 crashes (*8*, *9*).

Crashes involving animals, particularly wildlife, typically occur on rural two-lane roads and are often single-vehicle collisions (*10*). As a primary step to providing streamlined guidance and best-practice solutions to Louisiana, it is necessary to analyze the characteristics of crashes involving animals on the road, especially

[1]Department of Civil Engineering, University of Louisiana, Lafayette, LA
[2]Ingram School of Engineering, Texas State University, San Marcos, TX
[3]Department of Civil Engineering, Auburn University, Auburn, AL

*M. Ashifur Rahman, Julius Codjoe, Elisabeta Mitran are also affiliated to Louisiana Transportation Research Center, Baton Rouge, LA

Corresponding Author:
M. Ashifur Rahman, ashifur@outlook.com; ashifur@louisiana.edu; ashifur.rahman@la.gov

considering the dearth of previous research on the topic in this state.

To find the collective association of attributes from AVC data, which are categorical in nature, this study utilized multiple correspondence analysis (MCA) and association rule mining (ARM). The hierarchical clustering (HC) approach helps identify crash clusters on the two-dimensional approximation of the categorical dataset developed by MCA. The research team applied the a priori algorithm of the ARM approach based on some prespecified measures. Since most crashes caused by animals on the road do not result in visibly injured drivers or passengers, three separate assessments of crashes with varying degrees of severity were conducted. The study specifically analyzed three severity groups based on the KABCO injury severity scale (K—fatal injury, A—severe injury, B—moderate injury, C—possible/complaint injury, O—no injury): major injury (KA), minor injury (BC), and no injury (O) crashes.

Literature Review

Exploration of AVC Contributing Factors

Researchers have applied advanced statistical models to AVC data to identify the significance of influential factors. Murphy and Xia modeled the risk of AVCs on a segmented highway in Western Australia using a hierarchical Bayesian model involving multivariate Poisson lognormal regression (*11*). Farming on both sides of the road, a mixture of farming and roadside vegetation, and roadside vegetation each positively affect AVCs, whereas speed limits and horizontal curves negatively affect them. Stapleton et al. performed a cross-sectional analysis of deer-vehicle crashes (DVC) on two-lane rural highways using data from Michigan and mixed effects negative binomial regression models (*12*). DVC occurrence was significantly affected by speed-related factors such as lane width, shoulder width, horizontal curvature, and peak level of service. Wider lanes had a greater occurrence of DVCs, whereas horizontal curves with design speeds lower than the speed limit, along with roads that had wider shoulders, had fewer DVCs. Wilkins et al. used ordinary least-squares (OLS) regression analysis across Texas to study AVCs (*13*). The authors found that less densely populated rural counties and those with fewer vehicle-miles traveled (VMT) per capital but more lane-miles per capita tended to have the greatest number of AVCs per VMT. Large crossing structures at the highway link level may return benefit-to-cost ratios of nearly 3.0, whereas lower-cost counterparts deliver ratios up to 30.

Ashraf and Dey used Bayesian spatiotemporal models and 5 years of crash data to identify and prioritize DVC hotspots (*14*). Forest area, vegetation, and wetland percentages were positively associated with DVC frequency, whereas the percentage of developed land use was negatively associated with DVC frequency. DVCs are also affected by deer population, so deer population management is important to minimize DVC risks. Suburban areas with mixed land use conditions were also shown to have a higher risk of DVCs. Some researchers used machine learning models to explore AVC data. Moghaddam et al. proposed five machine learning-based prediction models for AVCs in the presence of categorical features, which were developed using eXtreme gradient boosting (XGBoost), logistic regression, CatBoost, random forest, and light gradient boosting machine (LGBM) (*15*). The CatBoost model had the highest accuracy (78.52%) and was subsequently the most suitable model for predicting AVCs.

AVC Injury Severity Studies

Injury severity analysis shows the impact of contributing factors on different crash injury levels. Al-Bdairi et al. investigated the determinants of driver injury severity in AVCs using data from Washington state from 2012 and 2016 and the mixed logit model, mixed logit model with heterogeneity in means, and mixed logit model with heterogeneity in means and variances (*16*). There were many parameters that can potentially increase the likelihood of KA AVCs, such as freeways/expressways, daylight crashes, early morning crashes, dry road surfaces, and clear weather conditions. In addition, the model fit can be improved by accounting for the heterogeneity in the means (and variances) of the random parameters.

Ahmed et al. studied DVCs and resulting injury severity using two random parameters binary logit models, a novel variant of ordered probability means, and data from 2018 from the Pennsylvania Department of Transportation (*17*). Rural locations, dark lighting conditions, and deer breeding season increase the likelihood of witnessing deer. Drivers were more likely to hit deer on roads with speed limits over 55 mph, during deer breeding season, and/or if the driver is female. Major injuries and fatalities are associated with airbag deployment and post-crash overturning of the vehicle, where the use of restraints was associated with preventing injuries or fatalities. Gharraie and Sacchi used structural equation modeling (SEM) with generalized (ordered probit) links to study the severity outcomes of wildlife-vehicle crashes (WVCs) in Canada (*18*). The authors looked at the driver's speeding attitude (SA), the driver's visibility impairment (VI), and crash severity. SA and VI were both shown to positively affect crash severity, although SA was the most influential factor. Road surface conditions strongly affected the SA measurement model, and weather conditions strongly affected VI.

Other Exploratory Investigative Approaches

It is important to understand the cost and legal perspectives of AVCs on human safety. Abra et al. explored this issue using data from Brazil (*19*). There were 2,611 AVCs per year (3.3% of total crashes), of which 18.5% resulted in human injuries/fatalities. The average cost of an AVC was R$ 21,656 (U.S. $9,629). The Brazilian legal system overwhelmingly holds road administrators liable for AVCs; road administrators spend R$ 2,463,380 (U.S. $1,005,051) per year compensating victims and are expected to keep wild and domestic species off the road. The authors suggest improved coordination between the laws relating to AVCs and human safety along with better management practices for abandoned domesticated animals. Cherry et al. aimed to describe conditions and circumstances involving AVCs using U.S. National Park Service (NPS) law enforcement motor vehicle crash (MVC) data to guide traffic and wildlife management for the prevention of AVCs in select NPS units (*20*). Northeast and intermountain NPS regions had the largest percentage of the total AVC burden. The fall showed the highest counts of AVCs for the national capital, northeast/southeast, and northeast regions. Winter showed the highest counts for the southeast region, and summer showed the highest counts for intermountain and Pacific West regions. AVCs in select NPS units were twice as high as the national average.

Some studies explored innovative topics associated with AVCs. Backs et al. measured train audibility and developed a physical model to simulate train audibilities using 10 locations in a mountain park (*21*). Both measured and simulated values were compared at locations of wildlife-train collisions in the past 35 years. Locations with lower measured train audibilities and the lowest quartile of simulated audibilities had more wildlife collisions. Train audibility was reduced by hill height in approaches around curves, but track curvature was not a good model to predict audibility overall. Adjacent road traffic background noise also reduced train audibility, along with high train speeds and down-grade travel. Conway et al. used a survey to study the physical and emotional impacts of moose vehicle collisions and DVCs and determined the actions of first responders/healthcare services accessed following the collision (*22*). Vehicle collisions with moose had more injuries and higher usage of healthcare resources and providers than deer. In addition, they required a greater response from first responders. The emotional impacts of moose and DVCs, however, were very similar; around half of the respondents in both groups felt anxious when reminded of the collision. Llagostera et al. developed a new algorithm-based approach to determine minimum paths between vertices in weighted networks to obtain the safest route between two points in a road structure based on WVC point patterns together with other road variables such as road speed limits, traffic volume (traffic flow information), and vegetation density around roads (*23*). They then demonstrated the use of this network using a real data set with the locations of 491 WVCs in north-east Spain.

A few researchers explored video data to understand the impact of AVCs. Rea et al. aimed to study the dynamics of moose-vehicle collisions using vehicle dash-mounted camera (dash cam) videos of moose-vehicle interactions posted on YouTube (*24*). The authors employed a parsimonious logistic regression model to assess the data. "Vehicle slows" was the only variable found to affect the difference between moose-vehicle collisions and near misses. Brieger et al. used 24,800 h of video data from 2,841 animal-vehicle encounters to classify animal behavior before and during vehicle contact (*25*). The authors analyzed the data using ordinal Bayesian mixed-effect regression models. Animal attentiveness, behavior a priori, access to cover, vehicle type, and biological seasonality were important predictors of animal responses to oncoming vehicles. Overall, collision risk for common European mammals depends on vehicle type, road layout, and species-specific behavioral repertoire (i.e., the attentiveness of the animal and behavioral state before a vehicle approaches). The authors also found that wildlife warning reflectors did not alter behavioral reactions and failed to reduce WVC risk.

The literature cited above shows that conducting an analysis of contributing factors for AVCs and associated injury severity risks by jurisdiction is the most credible approach toward the identification of implications for potential countermeasures. Before taking advantage of new innovative countermeasures to minimize AVCs, it is imperative to explore common patterns by taking into account a set of crash, roadway, and driver characteristics. This paper addresses this gap for Louisiana with the analysis of the latest AVC data available and examines the issue further by investigating the collective association of the possible contributing factors.

Data

AVC data were initially acquired from Louisiana DOTD's online password-protected library of crashes, "Crash1" (*26*). Then, additional crash characteristics data were obtained from Microsoft Access-formatted crash data files which are created annually and are more thorough. To reflect recent crashes caused by the presence of animals on roadways, AVCs that occurred from 2015 to 2020 were exported from the two crash databases and were then merged to create a complete dataset, resulting in a total of 14,349 crashes after data cleaning. The distribution of severity groups in the final dataset is as follows—KA: 48 (0.33%), BC: 2,198 (15.32%), and O: 12,103 (84.35%).

To facilitate an insightful analysis, a group of relevant variables was chosen, including, but not limited to,

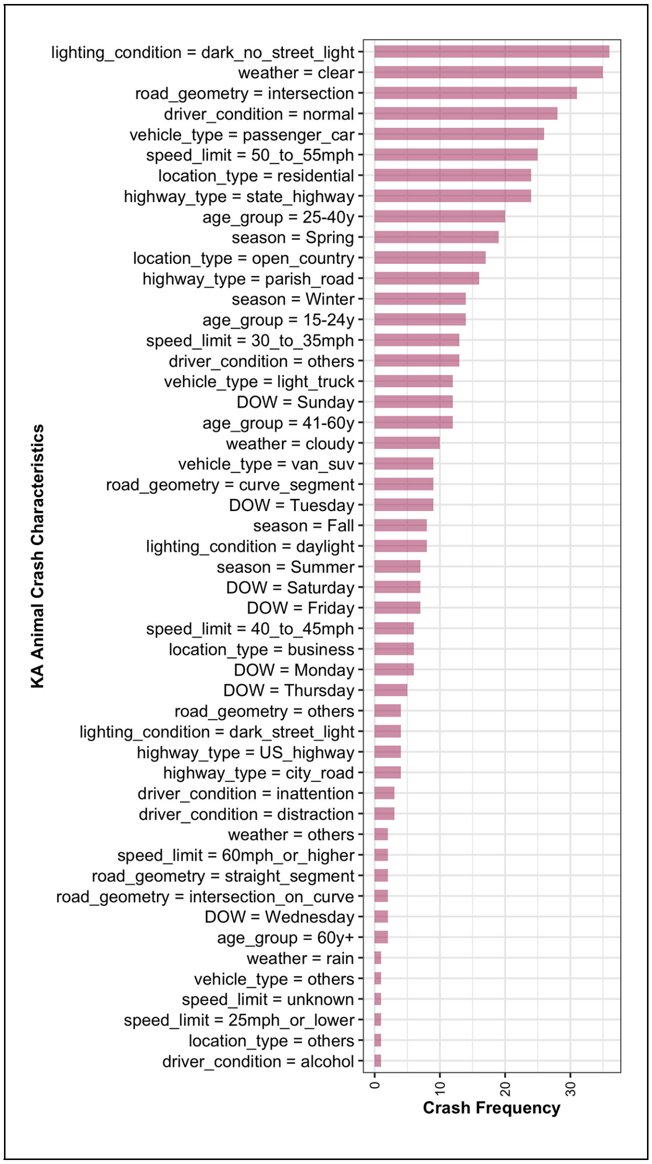

Figure 1. Descending-ordered frequencies of animal-vehicle crash (AVC) characteristics resulting in major injury (KA).
Note: DOW = day of week.

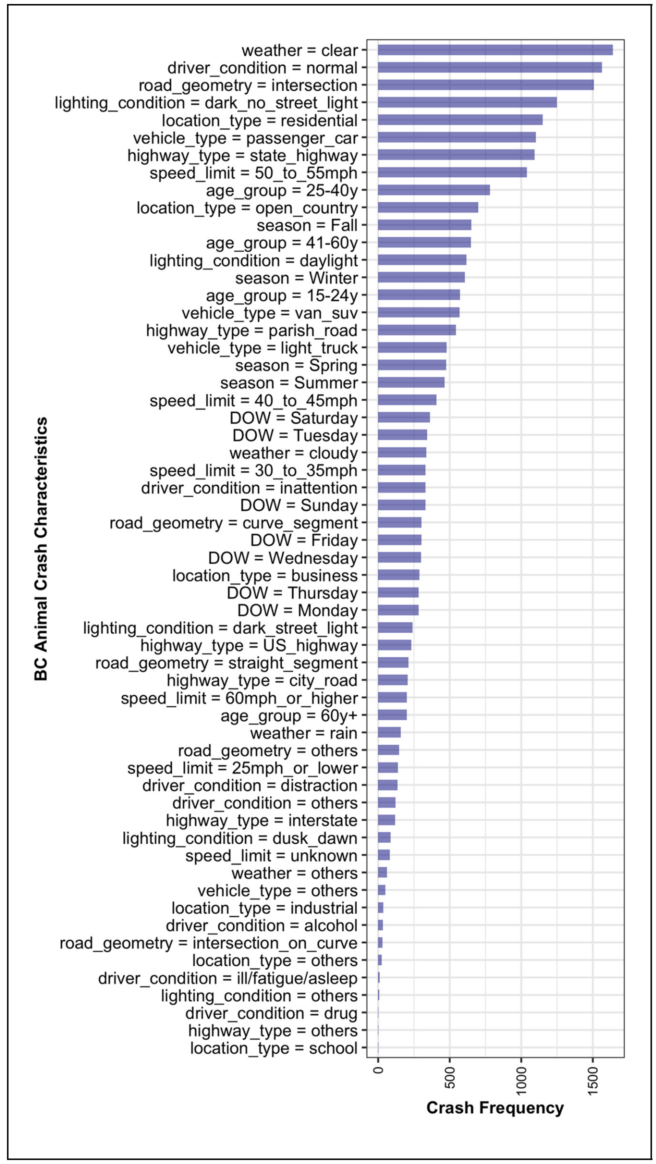

Figure 2. Descending-ordered frequencies of animal-vehicle crash (AVC) characteristics resulting in minor injury (BC).
Note: DOW = day of week.

roadway configuration and location characteristics, temporal characteristics such as season and day of the week, and driver-related factors such as driver age and driver condition. Before conducting the analysis, an overview of the variables was provided by examining their distribution, since the analysis was divided by severity groups. Figures 1 to 3 present the attributes of AVCs leading to KA, BC, and O severity in descending order of frequency, respectively, to emphasize the most common ones.

The leading 10 characteristics in the three severity groups exhibit considerable similarity. Notably, the absence of streetlights is prevalent across all three categories. Additionally, various typical driving conditions, such as clear weather, regular driving, and crashes with passenger cars, are among the top 10 attributes in all three groups. The fall season is proportionally more associated with BC and O crashes. With respect to location type, crashes primarily occur in open country and residential areas. In the case of the latter, it typically pertains to dispersed rural residential settlements.

Methodology

As standard statistical parametric models establish links based on the distributions of independent and dependent

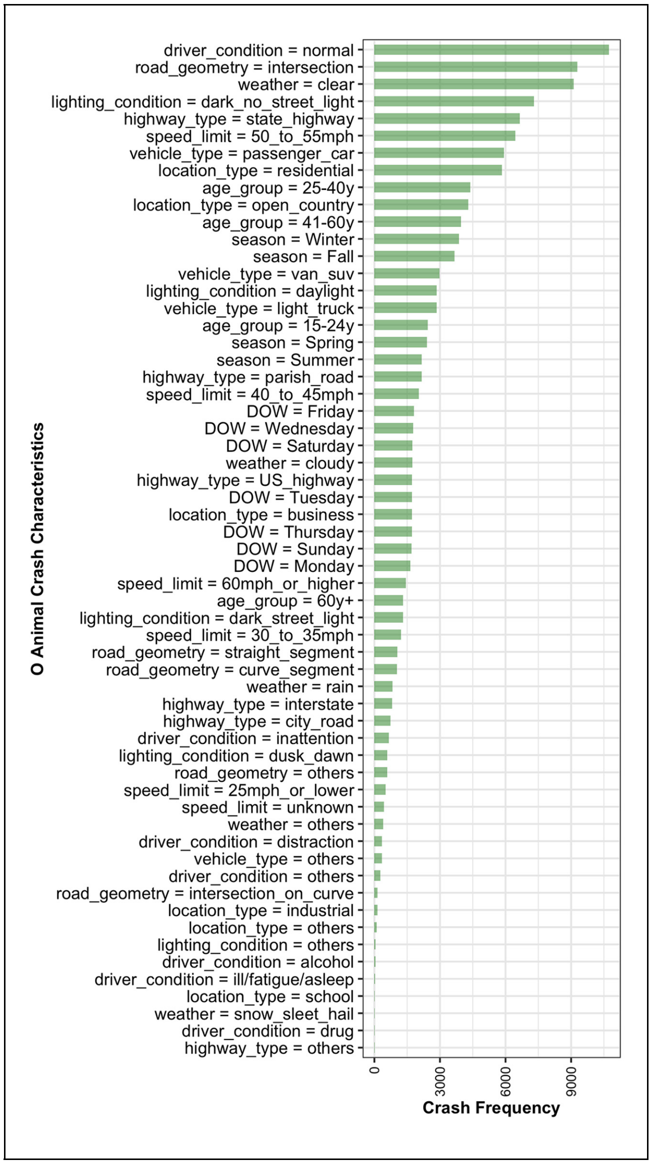

Figure 3. Descending-ordered frequencies of animal-vehicle crash (AVC) characteristics resulting in no injury (O).
Note: DOW = day of week.

variables, they have limited capacity for disentangling complex interactions between many variables in high-dimensional datasets. With an increase in the number of factors, these deficiencies worsen beacause of an increase in scatterings, leading to erroneous results. To overcome these issues, this study applied MCA and ARM. We utilized R software to conduct both analyses (*27*).

Multiple Correspondence Analysis (MCA)

MCA, a multivariate statistical method used to analyze relationships among categorical variables, is being increasingly used in transportation safety research. The main idea behind MCA is to find a set of new variables called "dimensions," which are linear combinations of the original variables that capture the maximum amount of information in the data. These dimensions are orthogonal to each other and can be plotted in a two- or three-dimensional space to visualize the patterns in the data. Dimensional reduction offered by MCA enables the identification of clusters with implicit relationships among attributes. Several transportation safety studies have used MCA to investigate a multitude of issues, such as vehicle-pedestrian crashes, wrong-way-driving crashes, run-off-road crashes, and so forth (*28–31*).

The data in MCA is organized in a contingency table, where rows correspond to observations and columns correspond to variables. The indicator matrix is a binary matrix that encodes the categories of each variable. The row and column profiles are computed by dividing the frequencies in the contingency table by the marginal frequencies. The singular value decomposition (SVD) is used to decompose the matrix of standardized residuals into a product of matrices that contain information about the row and column profiles. The row and column profiles are plotted in a low-dimensional space (usually 2D) where the distances between points reflect the degree of association between categories. For further details on the procedure, Das and Sun can be referred to (*28*).

Hierarchical Clustering (HC) on the MCA

The steps of HC in MCA results involve: 1) calculating the distances between the observations using a chosen distance metric, 2) creating a proximity matrix based on the calculated distances, and 3) grouping the observations with the shortest distance together to form a cluster. For further details on the HC in MCA, Husson et al. can be referred to (*32*).

Association Rule Mining (ARM)

ARM is a prominent technique in several disparate fields such as market basket analysis, product recommendation, and medical diagnostics for finding intriguing and nontrivial relationships between variables, and it has recently gained popularity in highway safety analysis (*33–37*). Transportation safety domains have been using ARM to explore underlying patterns. There is potential for road safety agencies to use ARM to find coexisting crash features for decision-making by focusing on specific crash issues. Association rules based on established support and confidence thresholds can express links between variables in a variety of contexts without limiting the type of variables (independent or dependent). In addition, ARM can handle both large and small datasets, unlike other non-parametric data mining techniques (*38*).

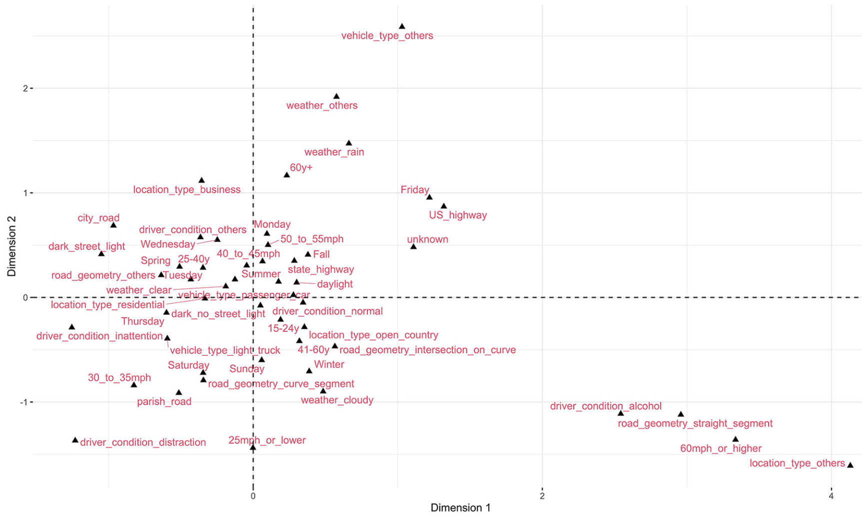

Figure 4. The Biplot of attributes from the animal-vehicle crashes (AVCs) with major injury (KA).

Theoretical Background of ARM. $T = \{t_1,\ t_2,\ t_3,\ \ldots,t_N\}$ indicates a dataset of N crashes in which each row comprises a set of crash attributes. An association rule is expressed in the form of $A \rightarrow B$. In this form, A is known as the antecedent and is presented on the left side, whereas B is the consequent which is presented on the right side. Three measures are typically used to quantify the quality of the rules: support, confidence, and lift. These measures can be expressed using the following equations:

$$\text{Support of the antecedent}, S(A) = \frac{n(A)}{N} \tag{1}$$

$$\text{Support of the consequent}, S(B) = \frac{n(B)}{N} \tag{2}$$

$$\text{Support of a rule}, S\ (A \rightarrow B) = \frac{n(A \cap B)}{N} \tag{3}$$

$$\text{Confidence of a rule}, C\ (A \rightarrow B) = \frac{S\ (A \rightarrow B)}{S\ (A)} \tag{4}$$

$$\text{Lift of a rule}, \mathit{Lift}\ (A \rightarrow B) = \frac{S\ (A \rightarrow B)}{S\ (A)\ \times\ S\ (B)} \tag{5}$$

where
N = count of crashes in the dataset,
$n(A)$ = count of occurrences with the antecedent A,
$n(B)$ = count of occurrences with consequent B, and
$n(A \cap B)$ = count of occurrences with the antecedent A and the consequent B together.

In association rule discovery, the support of a rule indicates the associative characteristics that occur most frequently. The confidence of an association rule represents the conditional probability of the consequent given the antecedent. The lift value represents the co-occurrence of the antecedent on the conditional likelihood of the consequence (*39*, *40*). A rule with a lift value of more than 1 implies a positive relationship between the antecedent and the consequent, whereas a value of less than 1 suggests a negative relationship.

Based on the format of a given dataset, researchers have created several efficient algorithms for identifying common itemset patterns (*38*). The a priori approach was designed for searching recurrent itemsets with given threshold values by assuming the subset of itemsets to be frequent. In the a priori process, frequent subsets are extended one item at a time to generate candidates, and groups of candidates are evaluated against the data using a "bottom-up" method. The algorithm ends when no additional successful expansions are discovered. A priori uses a breadth-first search to efficiently find candidate itemsets. More details of the apriori algorithm's specifications can be found in Agrawal and Srikant (*41*).

Results and Discussions

Results of MCA and Hierarchical Clusters

Separate analyses of MCA were conducted on three AVC datasets disaggregated by KA, BC, and O severity. Biplots of attributes resulting from MCA illustrated the projection of attributes on the major two dimensions, as presented in Figures 4 to 6 for KA, BC, and O crashes, respectively.

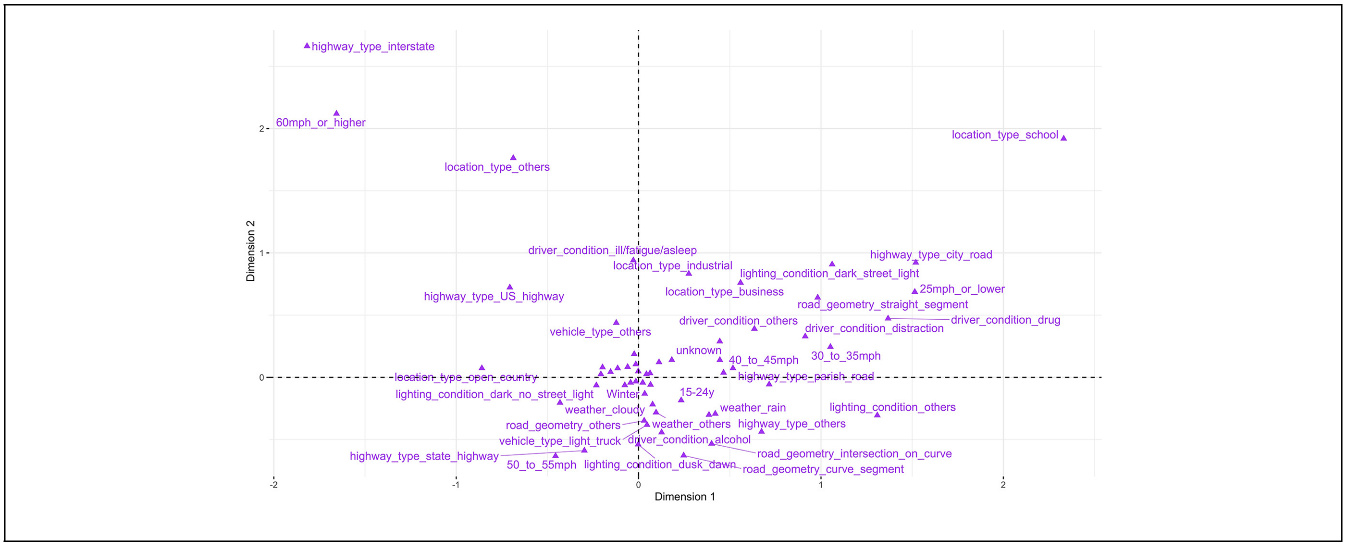

Figure 5. The Biplot of attributes from the animal-vehicle crashes (AVCs) with minor injury (BC).

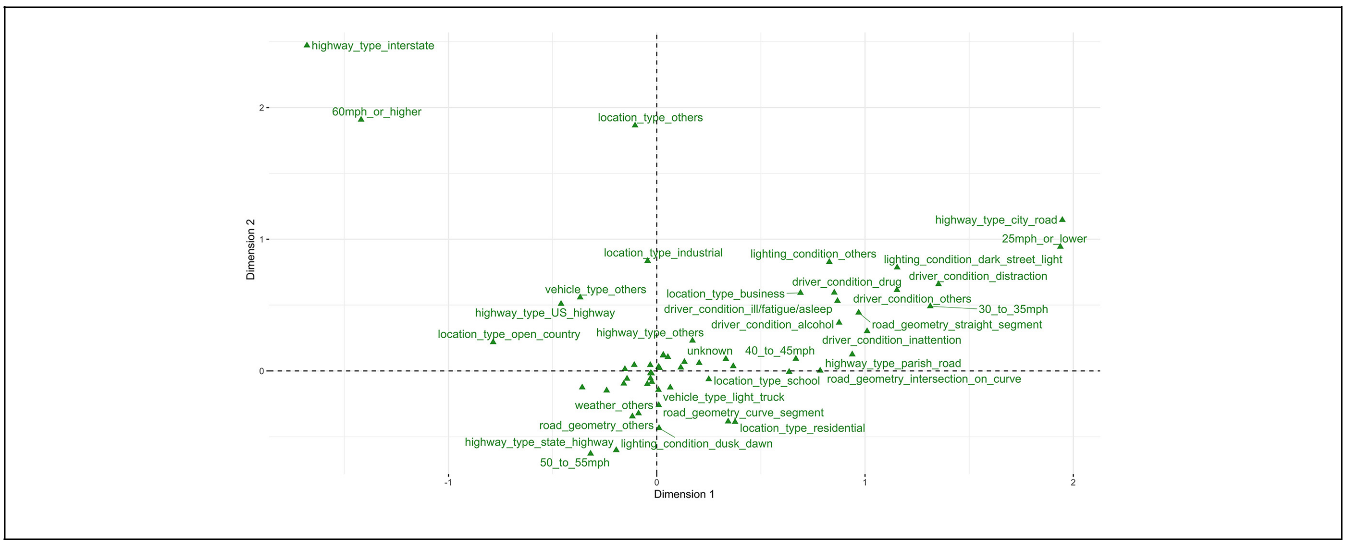

Figure 6. The Biplot of attributes from the animal-vehicle crashes (AVCs) with no injury (O).

MCA and HC Results for Major Injury (KA) Crashes. The interpretation of associative relationships between attributes, as represented on a biplot, depends on 1) the relative positioning of the remaining attributes under the same variable on a low-dimensional approximation, and 2) their individual proximities with other potentially related attributes. From the proximity of attributes in Figure 4 (the MCA biplot of KA AVCs), it can be deduced that parish roads (equivalent to county roads in other states) are more closely associated with a speed limit of 30–35 mph (lower left quadrant), while state highways are more closely associated with speed limits of 40–45 mph or 50–55 mph (upper right quadrant). This does not necessarily indicate that parish roads are not associated with 40–45 mph or 50–55 mph, but rather suggests that in relation to AVCs resulting in KA injuries, parish roads have a relatively higher association with 30–35 mph.

It should also be considered that low-dimensional approximation of attributes based on the projections of their data points does not necessarily suggest that they cannot be linked with multiple associations. Nevertheless, several prior studies have relied on the manual exploration of associated attributes based on their perceived proximity (*31*). As the HC algorithm was employed on the primary dimensional plane from the MCA of the KA dataset aiming to systematically find associations among attributes, 3 clusters were derived by connecting several potentially associated attributes.

Table 1. Hierarchical Clusters on Animal-Vehicle Crashes (AVCs) Resulting in Major Injury (KA)

Cluster	Cla/Mod	Mod/Cla	*p*-value	*v*-test
Cluster 1				
Highway type = parish road	93.75	60	<0.0001	4.137
Vehicle type = light truck	100	48	<0.0001	3.961
Road geometry = curve segment	100	36	0.001	3.235
Cluster 2				
Highway type = city road	100	20	0.025	2.243
Location type = business	83.333	25	0.042	2.037
Speed limit = 50–55 mph	56	70	0.042	2.029
Cluster 3				
Speed limit = 60 mph or higher	100	66.667	0.003	3.005
Road geometry = straight segment	100	66.667	0.003	3.005
Season = winter	21.429	100	0.021	2.307

Note: Cla/Mod = distribution of significant attributes across clusters; Mod/Cla = distribution within-cluster.

Table 1 presents the hierarchical clusters on the KA crashes obtained from its MCA. A v-test value higher than 1.96 relates to a p-value of less than 0.05. A positive sign of the v-test indicates that the mean of the cluster is over-expressed for the attribute, and similarly, a negative sign implies the cluster is under-expressed for the attribute. For the purpose of reasonable meaningful interpretation, attributes with positive v-test values as well as p-values of less than 0.05 are presented for each cluster. As the positive v-test indicates the mean of the cluster is greater than the overall mean, alternatively higher v-test values for an attribute imply a stronger association with the cluster. As positive v-test values comparatively decrease in magnitude, it can be inferred that the strength of the association weakens. Therefore, top attributes ordered by the higher to lower values of the v-test are more likely to be associated together.

The attributes within a cluster can be presented through their association with categories estimated by several measures. Distributions of significant attributes across clusters (Cla/Mod), distributions within-cluster (Mod/Cla), and associated p-values and v-test values are presented in Table 1. For example, 93.75% parish road KA crashes belong to cluster 1, and 60% of all KA crashes of cluster 1 belong to parish road.

Cluster 1 implicates a collective association of AVC with KA severity by connecting parish road, light trucks, and curved road segments. All of these attributes are located on the lower left quadrant of the biplot in Figure 4. Cluster 2 suggests a potential association between crashes involving killed animals and the attributes of city roads, business locations, and speed limits of 50–55 mph. Cluster 3 is associated with higher speed limits of 60 mph or more, straight road segments, and the winter season.

MCA and HC Results for Minor Injury (BC) Crashes. Figure 5 displays the MCA results of the BC crashes in a biplot, which provides a two-dimensional representation of the attributes. The BC dataset is large, resulting in complex interrelationships among its attributes compared with smaller datasets such as KA crashes. The dispersed nature of the attributes in the biplot indicates that it is practical to extract their collective association by using HC.

Table 2 presents the results of HC applied to the BC dataset of AVCs, which produced five distinct clusters. For the sake of simplicity and brevity, only v-test values are shown in the table. Each cluster is composed of multiple attributes, and they are presented based on variables for better comparison of their composition. Cluster 1 stands out with high v-test values and high across-cluster percentages (i.e., Cla/Mod values), and it associates AVCs resulting in BC with open country locations, interstate highways, and speed limits of 60 mph or higher. This cluster also shows associations with U.S. highways, intersections, and dark streetlight conditions. The concentration of crashes in this cluster occurs in open country areas that are particularly susceptible to AVCs involving injury, especially when animals are crossing high-speed roads with high traffic volumes. Additionally, the absence of lighting may contribute to reduced visibility and increase the risk of collisions.

The crashes in cluster 2, which are also somewhat concentrated in open country locations, show an apparent association with state highways and a speed limit of 50–55 mph, which are likely interrelated. This cluster also shows notable associations with curve segments and the absence of streetlights. This may be because of reduced visibility and reduced driver reaction time on curves.

Cluster 3 is associated with business locations, U.S. highways, speed limits of 40–45 mph, inattentive drivers, and drivers aged over 40 years. This cluster is further associated with daylight, the summer season, and Wednesdays.

Table 2. Hierarchical Clusters on Animal-Vehicle Crashes (AVCs) Resulting in Minor Injury (BC)

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test
Location type															
Open country	20	72	12.036	62	45	11.461	–	–	–	–	–	–	–	–	–
Residential	–	–	–	–	–	–	–	–	–	33	79	13.681	14	66	4.689
Business	–	–	–	–	–	–	64	61	22.188	–	–	–	21	24	4.984
Industrial	–	–	–	–	–	–	–	–	–	–	–	–	33	5	3.499
Highway type															
Interstate	94	58	23.788	–	–	–	–	–	–	–	–	–	–	–	–
U.S. highway	28	34	9.415	–	–	–	40	31	10.761	–	–	–	–	–	–
State highway	–	–	–	79	89	34.433	–	–	–	–	–	–	–	–	–
Parish road	–	–	–	–	–	–	–	–	–	82	92	37.808	–	–	–
City road	–	–	–	–	–	–	–	–	–	–	–	–	74	61	23.417
Speed limit															
25 mph or lower	–	–	–	–	–	–	–	–	–	–	–	–	72	40	17.753
30–35 mph	–	–	–	–	–	–	–	–	–	70	48	20.733	23	30	6.4
40–45 mph	–	–	–	–	–	–	28	38	8.749	38	32	8.099	–	–	–
50–55 mph	–	–	–	82	88	36.137	–	–	–	–	–	–	–	–	–
60 mph or higher	93	97	33.434	–	–	–	–	–	–	–	–	–	–	–	–
Road geometry															
Straight segment	–	–	–	–	–	–	34	24	8.054	–	–	–	33	27	8.807
Curve segment	–	–	–	53	17	3.327	–	–	–	31	19	3.987	–	–	–
Intersection	12	90	7.364	–	–	–	–	–	–	–	–	–	–	–	–
Lighting condition															
Daylight	–	–	–	–	–	–	30	62	13.41	27	34	3.02	–	–	–
Dark (streetlight)	–	–	–	–	–	–	–	–	–	–	–	–	55	52	18.121
Dark (no streetlight)	12	75	5.518	59	76	16.339	–	–	–	–	–	–	–	–	–
Weather															
Clear	10	83	2.839	–	–	–	–	–	–	–	–	–	–	–	–
Cloudy	–	–	–	–	–	–	18	20	2.533	–	–	–	–	–	–
Rain	–	–	–	–	–	–	–	–	–	–	–	–	20	13	3.311
Vehicle type															
Passenger car	–	–	–	47	54	3.027	–	–	–	–	–	–	15	68	5.951
Van/SUV	11	32	2.047	–	–	–	22	42	6.454	–	–	–	–	–	–
Light truck	–	–	–	48	24	2.221	–	–	–	31	30	4.975	–	–	–
Age group (years)															
25–40	–	–	–	–	–	–	–	–	–	–	–	–	14	43	2.741
15–24	–	–	–	–	–	–	–	–	–	26	30	2.456	14	33	2.516
41–60	–	–	–	–	–	–	18	38	3.431	–	–	–	–	–	–
60 +	–	–	–	–	–	–	34	22	7.63	–	–	–	–	–	–

(continued)

Table 2. (continued)

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
	Cla/Mod	Mod/Cla	v-test	Cla/Mod	Mod/Cla	v-test	Cla/Mod	Mod/Cla	v-test	Cla/Mod	Mod/Cla	v-test	Cla/Mod	Mod/Cla	v-test
Driver condition															
Normal	10	83	3.91	50	81	9.249	–	–	–	–	–	–	–	–	–
Distraction	–	–	–	–	–	–	–	–	–	–	–	–	37	20	8.039
Inattention	–	–	–	–	–	–	28	31	7.616	–	–	–	–	–	–
Ill/fatigue/asleep	–	–	–	–	–	–	–	–	–	–	–	–	40	2	2.276
Season															
Spring	–	–	–	–	–	–	–	–	–	28	27	3.338	–	–	–
Summer	–	–	–	–	–	–	23	35	5.996	–	–	–	–	–	–
Day of week															
Wednesday	–	–	–	–	–	–	19	19	2.628	–	–	–	–	–	–

Note: Cla/Mod = distribution of significant attributes across clusters; Mod/Cla = distribution within-cluster; SUV = sports utility vehicle. Long dashes indicate significant positive values don't exist for the attribute.

The crashes in cluster 4 appear to be more strongly associated with parish roads and speed limits of 40–45 mph, with other notable associations including residential areas, curve segments, spring season, and light trucks. Despite the prevalence of AVCs during the winter and fall seasons because of increased animal movements in Louisiana, AVCs with BC severity in summer and spring have been partially captured by cluster 3 and cluster 4, respectively.

Cluster 5 appears to be largely associated with low speeds (25 mph or lower), city roads, and darkness with the presence of lighting. Although AVCs may appear to be unusual in business areas (cluster 3) and on city roads (cluster 5), it must be noted that Louisiana is a rural state and often has business activities near jungles or swamps. Crashes may occur either by colliding with or avoiding animals such as alligators, raccoons, squirrels, and so forth in these areas, leading to driver or occupant injuries.

MCA and HC Results for No Injury (O) Crashes. The biplot of O crashes presented in Figure 6 illustrates the two-dimensional approximation of its attributes from MCA results. The results of HC are presented in Table 3.

Five clusters were also derived from the O AVCs, which typically involve damaged or disabled vehicles and other property damages. Very high v-test values in this table (noted as "Inf") indicate a p-value that tends to be very close to zero. Similar to BC crashes, cluster 1 in O crashes is also strongly associated with open country locations, interstate highways, and a speed limit of 60 mph and higher. The association with dark with no streetlight, state highways, and intersections is also similar to cluster 1 of BC crashes. Unlike BC crashes, this cluster in O crashes appears to be associated with the fall season.

Cluster 2 in O crashes is also associated with state highways and a 50–55 mph speed limit; however, it seems to be concentrated in open country and residential areas. This cluster can be further associated with the absence of streetlights, intersections, middle-aged drivers (41–60 years), winter season, and so forth. Cluster 3 of O crashes shows associations of drivers driving in normal conditions, on straight segments, with speed limits of 40–45 mph, on U.S. highways, within the 25–40-year-old age group, on Sunday, and in both fall and winter seasons. One study identified a higher proportion of AVCs during the weekend because of hunting (*44*).

Clusters 4 and 5 are primarily associated with crashes on parish roads and city roads, respectively, with speed limits ranging from 30–35 mph, 40–45 mph, and 50–55 mph. Cluster 5 appears to be significantly associated with city roads but significantly concentrated in business areas, unlike BC cluster 5 where crashes on city roads are distributed in residential, business, and industrial areas.

Table 3. Hierarchical Clusters on Animal-Vehicle Crashes (AVCs) Resulting in No Injury (O)

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test
Location type															
Open country	24	80	35.165	62	43	18.536	–	–	–	–	–	–	–	–	–
Residential	–	–	–	57	54	12.069	–	–	–	29	76	29.764	–	–	–
Business	–	–	–	–	–	–	63	69	Inf	–	–	–	22	47	23.502
Industrial	21	2	3.524	–	–	–	28	2	4.604	–	–	–	–	–	–
Highway type															
Interstate	97	61	Inf	–	–	–	–	–	–	–	–	–	–	–	–
U.S. highway	24	32	17.665	–	–	–	32	35	22.18	–	–	–	–	–	–
State highway	–	–	–	80	86	Inf	–	–	–	–	–	–	–	–	–
Parish road	–	–	–	–	–	–	–	–	–	87	85	Inf	–	–	–
City road	–	–	–	–	–	–	–	–	–	–	–	–	75	68	Inf
Speed limit															
25 mph or lower	–	–	–	–	–	–	–	–	–	43	10	13.381	55	34	30.374
30–35 mph	–	–	–	–	–	–	–	–	–	71	39	Inf	22	32	18.264
40–45 mph	–	–	–	–	–	–	32	41	25.072	34	32	19.382	10	24	5.312
50–55 mph	–	–	–	85	88	Inf	–	–	–	–	–	–	–	–	–
60 mph or higher	88	96	Inf	–	–	–	–	–	–	–	–	–	–	–	–
Road geometry															
Straight segment	–	–	–	–	–	–	26	18	12.352	–	–	–	30	40	25.344
Curve segment	–	–	–	–	–	–	–	–	–	34	16	12.651	–	–	–
Intersection	13	89	12.45	53	80	8.956	–	–	–	–	–	–	–	–	–
Lighting condition															
Daylight	–	–	–	–	–	–	–	–	–	20	26	2.822	13	46	14.342
Dark (streetlight)	–	–	–	–	–	–	48	40	33.816	–	–	–	26	42	24.013
Dark (no streetlight)	13	75	11.634	60	71	23.993	–	–	–	–	–	–	–	–	–
Weather															
Clear	11	80	3.886	–	–	–	14	80	4.684	–	–	–	–	–	–
Cloudy	–	–	–	57	16	5.148	–	–	–	–	–	–	–	–	–
Rain	–	–	–	–	–	–	17	9	3.501	–	–	–	–	–	–
Vehicle type															
Passenger car	12	54	3.54	–	–	–	16	59	8.376	–	–	–	7	53	2.077
Van/SUV	–	–	–	54	26	3.616	–	–	–	–	–	–	8	28	2.519
Light truck	–	–	–	56	26	6.295	–	–	–	22	28	5.297	–	–	–
Age group (years)															
25–40	–	–	–	–	–	–	14	40	3.051	–	–	–	–	–	–
15–24	–	–	–	–	–	–	–	–	–	27	30	12.475	–	–	–
41–60	–	–	–	57	36	8.913	–	–	–	–	–	–	–	–	–
60 +	–	–	–	–	–	–	–	–	–	–	–	–	13	21	8.407
Driver condition															
Normal	12	95	7.65	55	95	22.219	14	93	5.429	–	–	–	–	–	–
Distraction	–	–	–	–	–	–	–	–	–	–	–	–	37	20	8.039

(continued)

Table 3. (continued)

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test	Cla/Mod	Mod/Cla	*v*-test
Inattention	–	–	–	–	–	–	–	–	–	31	9	8.317	23	19	14.125
Ill/fatigue/asleep	–	–	–	–	–	–	–	–	–	46	—	3.378	18	—	1.999
Season															
Fall	13	35	3.813	–	–	–	14	33	2.819	–	–	–	–	–	–
Winter	–	–	–	53	33	2.704	14	34	1.964	–	–	–	–	–	–
Spring	–	–	–	–	–	–	–	–	–	–	–	–	10	28	5.944
Summer	–	–	–	–	–	–	–	–	–	20	20	2.711	8	22	3.185
Day of week															
Tuesday	–	–	–	55	15	3.433	–	–	–	–	–	–	–	–	–
Wednesday	–	–	–	55	16	3.354	–	–	–	–	–	–	–	–	–
Saturday	–	–	–	–	–	–	–	–	–	21	16	2.633	–	–	–
Sunday	–	–	–	–	–	–	15	17	3.218	21	16	2.897	–	–	–

Note: Cla/Mod = distribution of significant attributes across clusters; Inf = very high *v*-test value; Mod/Cla = distribution within-cluster; SUV = sports utility vehicle. Long dashes indicate significant positive values don't exist for the attribute.

Similar to BC cluster 5, O cluster 5 is also largely associated with low speed (25 mph or lower), dark with the presence of lighting, and so forth. In contrast to the BC clusters where crashes on curves were distributed in clusters 2 and 4, mainly representing state highways and parish roads, there was a significant association of O crashes with cluster 4, indicating crashes occurred mainly on parish roads. Other significant associations in cluster 4 are young drivers, inattentive and drowsy driving, and so forth. Young drivers have previously been found to be associated with KA crashes with animals (*10*). Drowsy driving has been associated with cluster 5, typically on city roads, which could largely be attributed to intoxicant use, considering Louisiana's more-than-national-average drinking habit (*45*). After examining the distributions across clusters (Cla/Mod) and within-cluster distributions (Mod/Cla) for age groups and driver conditions in both BC clusters and O clusters, it is difficult to attribute negative driving behaviors such as distracted, inattentive, or drowsy driving to specific age groups. This could be because of the recent expansion of such behaviors across all age groups (*46*).

ARM Results

While MCA was used to derive the clusters of three segregated injury severity groups of AVCs, ARM was utilized in identifying key associative characteristics for each severity group on an aggregate level. Because of the uneven distribution of severity groups in the dataset, a minimum threshold for support and confidence was set based on each group's proportion in the whole dataset. By specifying a severity group in the consequent, two key characteristics were identified regardless of the specific clusters. The results were ordered by support value to identify the most frequent associative characteristics. Tables 4 to 6 present the top 10 associative characteristics for KA, BC, and O crashes from a total of 74, 3,189, and 9,421 rules, respectively.

ARM Results of Major Injury (KA) Crashes. Because of the smaller size of the KA dataset, MCA provides little context; however, the ARM provides more details in this context. The factors include the presence of a residential area and a parish road, inattention by the driver in clear weather, and speed limits of 30–35 mph (see Table 4). The study also found that a residential area with speed limits of 30–35 mph and dark lighting conditions on a parish road, as well as a curve segment on a residential area, are significant factors in such crashes in reference to the complete AVC dataset. Furthermore, the type of road, whether it is a parish road or a city road, is also a significant factor. The top association rule identified in the analysis was the combination of a dark lighting

Table 4. Top 10 Association Rule Mining (ARM) Rules of Major Injury (KA) Crashes

Antecedent	Support (%)	Confidence (%)	Lift
(Lighting condition = dark [no streetlight]) + (highway type = parish road)	0.1	1	3.003
(Lighting condition = dark [no streetlight]) + (season = spring)	0.1	0.87	2.59
(Vehicle type = passenger car) + (lighting condition = dark [no streetlight]) + (age group = 15–24 years)	0.09	1.12	3.347
(Lighting condition = dark [no streetlight]) + (age group = 15–24 years)	0.09	0.7	2.088
(Weather = clear) + (lighting condition = dark [no streetlight]) + (season = spring)	0.08	0.89	2.65
(Lighting condition = dark [no streetlight]) + (highway type = parish road) + (speed limit = 30–35 mph)	0.07	1.83	5.465
(Location type = residential) + (speed limit = 30–35 mph)	0.07	1	3.001
(Speed limit = 30–35 mph) + (season = spring)	0.07	0.68	2.031
(Weather = clear) + (day of week = Sunday)	0.07	0.66	1.96
(Location type = residential) + (highway type = parish road) + (speed limit = 30–35 mph)	0.06	1.26	3.763

Table 5. Top 10 Association Rule Mining (ARM) Rules of Minor Injury (BC) Crashes

Antecedent	Support (%)	Confidence (%)	Lift
(Location type = residential) + (highway type = parish road)	2.77	20.66	1.349
(Location type = residential) + (age group = 15–24 years)	2.33	21.19	1.383
(Weather = clear) + (driver condition = inattention)	1.78	35.12	2.292
(Weather = clear) + (speed limit = 30–35 mph)	1.71	21.29	1.39
(Location type = residential) + (speed limit = 30–35 mph)	1.61	23.19	1.514
(Location type = residential) + (lighting condition = dark [no streetlight]) + (highway type = parish road)	1.52	21.14	1.38
(Highway type = parish road) + (speed limit = 30–35 mph)	1.48	21.92	1.431
(Highway type = parish road) + (age group = 25–40 years)	1.45	20.8	1.358
(Road geometry = curve segment) + (location type = residential)	1.27	24.36	1.591
(Location type = residential) + (highway type = city road)	0.94	24.5	1.599

condition (no streetlights) and the highway type being a parish road, which suggests that KA AVCs are more likely to occur in dark conditions on parish roads. Young drivers are more likely to be involved in KA AVCs at night. The rules also suggest KA crashes are more likely to occur in the spring season under dark lighting conditions, and young drivers are more prone to KA AVCs at night. The analysis also found that KA AVCs are more likely to occur in residential areas with a speed limit of 30–35 mph and during the spring season.

ARM Results on Minor Injury (BC) Crashes. The association rules identified from analyzing crash characteristics of AVCs resulting in minor injury to vehicle drivers or passengers suggest that certain factors may increase the likelihood of such crashes (see Table 5). The most significant association rule is the presence of a residential area and a parish road, indicating that such crashes are more likely to occur in residential areas on parish roads. The age group of 15–24 years is also found to have a significant association with crashes in residential areas. Weather and driver conditions are also found to be significant factors. Clear weather and inattention are also a top combination of factors involved in AVCs resulting in minor injury. Speed limits of 30–35 mph are also a significant factor when with clear weather conditions. The presence of a residential area with speed limits of 30–35 mph and dark lighting conditions on a parish road is also a significant factor. The presence of a curve segment in a residential area is more prone to AVCs resulting in minor injury. City roads and residential areas are another topmost frequent association.

ARM Results of No Injury (O) Crashes. The association rules show that several factors are associated with O AVC,

Table 6. Association Rule Mining (ARM) Results of No Injury (O) Crashes

Antecedent	Support (%)	Confidence (%)	Lift
(Road geometry = intersection) + (lighting condition = dark [no streetlight])	40.54	86	1.02
(Highway type = state highway) + (speed limit = 50–55 mph)	35.84	85.9	1.018
(Road geometry = intersection) + (highway type = state highway)	35.74	87.12	1.033
(Road geometry = intersection) + (speed limit = 50–55 mph)	35.45	87.24	1.034
(Weather = clear) + (driver condition = normal) + (speed limit = 50–55 mph)	31.39	88.21	1.046
(Lighting condition = dark [no streetlight]) + (speed limit = 50–55 mph)	30	85.74	1.017
(Road geometry = intersection) + (highway type = state highway) + (speed limit = 50–55 mph)	28.27	87.38	1.036
(Driver condition = normal) + (location type = open country)	27.77	87.73	1.04
(Driver condition = normal) + (age group = 25–40 years)	27.06	87.22	1.034
(Location type = open country) + (lighting condition = dark [no streetlight])	21.79	86.07	1.02

including lighting conditions, speed limit, road geometry, driver condition, and location type (see Table 6). However, some of these factors are more strongly associated with crashes than others, indicating that certain combinations of factors may have a greater impact on the occurrence of these types of crashes. For example, crashes that occurred at intersections in dark conditions with no streetlights are the relatively most frequent crashes. Similarly, crashes on state highways with a speed limit of 50–55 mph is another frequent group of characteristics. The occurrence of crashes at intersections on state highways with a speed limit of 50–55 mph indicates a more frequent association for O crashes. Other notable associations are clear weather, normal driver condition, and speed limits of 50–55 mph, dark conditions with no streetlights and a speed limit of 50–55 mph, and open country with dark conditions and no streetlights. It is noticeable that the lift values in O crashes are lower than KA and BC crashes, which suggests these O crash associative attributes may occur frequently but are not strongly associated with injury severity in comparison with the two prior injury groups.

Conclusions

AVCs have become an increasingly concerning issue in the U.S., and it is important to direct attention and research efforts toward them. However, given the potential variability in animal type, movement, and crash patterns across different states, conducting state-specific studies can be more reasonable and effective in identifying relevant factors and developing targeted strategies to mitigate the occurrence and severity of AVCs in each state. While it may be possible to apply countermeasures learned from other states to address AVCs, conducting an analysis to identify specific crash patterns can greatly assist in developing effective countermeasure strategies tailored to the local context.

Two data mining methods were applied in this study. The hierarchical clusters on the two-dimensional approximations of KA, BC, and O crash datasets imply that segregated severity crash patterns can largely be identified by highway type, corresponding location type, and speed limit characteristics. A substantially small dataset of 48 KA crashes provided three small clusters that represent parish roads, city roads, and high-speed roads with speed limits of 60 mph and higher. Crashes on high-speed roads have been linked to speeding and the concentration of wildlife in nearby areas (*10*).

The first two MCA clusters of BC crashes are on interstates, U.S., and state highways, and the key issues are related to visibility. A conventional countermeasure tailored to improving nighttime visibility on a high-speed roadway to lower the risk of AVCs is the placement of retroreflective warning signs in advance for locations with high animal activities; however, the effectiveness of conventional signs has been deemed unsatisfactory (*10*). Although more effective countermeasures are in practice to avoid the hindrance of visibility regardless of poor weather or nighttime such as radar-based animal detection systems, these countermeasures require a thorough investigation of animal activities, especially near high-speed roadways, for finding hotspots of animal crashes resulting from poor visibility (*47*).

Louisiana AVCs have a flexible distribution across seasons, and MCA reveals crashes in spring may be more concentrated in the area around parish roads for both BC and O crashes. The O crashes in fall and winter, which coincide with deer breeding season, may be more associated with clusters representing interstate, U.S., and state highways with speed limits of 50 mph or higher. Advance notification of known increased deer activity locations during fall and winter enabled through temporary roadside variable-message signs (VMS) have been proven to be economically beneficial. In Montana, using VMS as animal advisory messages has been found to be effective in reducing speed (*48*).

Association rules identified based on severity groups provide valuable insights into frequently overlooked but significant associations. ARM analysis revealed that, on high-speed state highways (50–55 mph), the most frequent association is with O crashes, suggesting that the presence of lighting on these roads should be investigated for potential effectiveness. Moderate speed parish roads (30–35 mph) are frequently associated with KA and BC crashes, particularly in residential areas and during the spring season, and often involve young drivers. Addressing speeding issues on these roads could help lower the risk of AVCs. Additionally, KA crashes involving animals tend to occur on curved roads in rural residential areas, which could potentially benefit from VMS facilities, especially if animal crossing activities are common.

Some limitations of the study need to be mentioned. There was a 63% drop in AVCs from 2019 to 2020, understandably linked to the COVID-19 pandemic. To avoid any bias, the research team included 6 years (2015–2020) of the latest data instead of 5 years. Future studies can investigate the change in temporal trends of AVCs because of COVID-19. It should also be emphasized that the occurrence of AVCs is influenced by the spatial arrangement and characteristics of wildlife as well as the connectivity of the urban network. Therefore, the findings and recommendations of this study must be interpreted with caution as they may oversimplify the risk of AVCs by not considering important factors such as the concentration of different animal types and their movements during peak breeding seasons. To accurately identify the animal species and roadway features that are most commonly associated with AVCs in Louisiana, it is crucial for researchers to take into account these spatiotemporal factors. To effectively reduce the risk of AVCs, a multidisciplinary approach that involves the Department of Transportation and Development and the Department of Wildlife and Fisheries should be taken to develop targeted interventions and mitigation strategies.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: M. Rahman, S. Das; data collection: M. Rahman, K. Abedi, M. Hossain; analysis and interpretation of results: M. Rahman, S. Das; draft manuscript preparation: M. Rahman, S. Das, J. Codjoe, E. Mitran, X. Sun. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

M. Ashifur Rahman https://orcid.org/0000-0001-6940-1599
Subasish Das https://orcid.org/0000-0002-1671-2753
Julius Codjoe https://orcid.org/0000-0003-1958-8695
Elisabeta Mitran https://orcid.org/0000-0002-1778-9832
Xiaoduan Sun https://orcid.org/0000-0001-7282-1340
Kwabena Abedi https://orcid.org/0000-0002-4071-6367
Md Mahmud Hossain https://orcid.org/0000-0002-2737-6951

References

1. IIHS-HLDI. Fatality Facts 2020: Collisions with Fixed Objects and Animals. *IIHS-HLDI Crash Testing and Highway Safety*. https://www.iihs.org/topics/fatality-statistics/detail/collisions-with-fixed-objects-and-animals. Accessed February 28, 2023.
2. LaChance, D. Collisions with Animals Rose 7.2% During the Pandemic, State Farm Analysis Says. *Repairer Driven News*, May 17, 2022.
3. Ang, J. Y., B. Gabbe, P. Cameron, and B. Beck. Animal–Vehicle Collisions in Victoria, Australia: An Under-Recognised Cause of Road Traffic Crashes. *Emergency Medicine Australasia*, Vol. 31, No. 5, 2019, pp. 851–855. https://doi.org/10.1111/1742-6723.13361.
4. IIHS-HLDI. Deer Season Means Animal Strike Season on the Roads. *IIHS-HLDI Crash Testing and Highway Safety*. https://www.iihs.org/news/detail/deer-season-means-animal-strike-season-on-the-roads. Accessed February 28, 2023.
5. Kinchen, D. M. Wildlife Crossings Provide Means to Prevent Thousands of Wildlife-Vehicle Collisions per Year in Louisiana. *Environmental Law Society*, February 4, 2018.
6. KPLC 7 News. Fatal Accident Involving Cow on Road Sheds Light on Livestock Owners' Responsibilities. *KPLC Virtual and VHF Digital Channel 7*. https://www.kplctv.com/story/26700798/fatal-accident-involving-cow-on-road-sheds-light-on-livestock-owners-responsibilities/. Accessed February 28, 2023.
7. LWLF. Estimated Deer Breeding Periods. *Louisiana Department of Wildlife and Fisheries*. https://www.wlf.louisiana.gov/page/deer-breeding-periods. Accessed March 5, 2023.
8. EverQuote. Largest Auto Insurance Companies in Louisiana. *EverQuote*. https://www.everquote.com. Accessed February 28, 2023.
9. StateFarm. Where Are Animal (Deer) Collisions Most Likely? *StateFarm: Simple Insights*. https://www.statefarm.com/simple-insights/auto-and-vehicles/how-likely-are-you-to-have-an-animal-collision. Accessed February 28, 2023.
10. Huijser, M. P., P. McGowan, A. Hardy, A. Kociolek, A. Clevenger, D. Smith, and R. Ament. *Wildlife-Vehicle Collision Reduction Study: Report to Congress*. Publication FHWA-HRT-08-034. Federal Highway Administration, Washington, D.C., 2017.

11. Murphy, A., and J. (Cecilia) Xia. Risk Analysis of Animal–Vehicle Crashes: A Hierarchical Bayesian Approach to Spatial Modelling. *International Journal of Crashworthiness*, Vol. 21, No. 6, 2016, pp. 614–626. https://doi.org/10.1080/13588265.2016.1209823.
12. Stapleton, S. Y., A. Ingle, and T. J. Gates. Factors Contributing to Deer–Vehicle Crashes on Rural Two-Lane Roadways in Michigan. *Transportation Research Record: Journal of the Transportation Research Board*, 2019. 2673: 214–224.
13. Wilkins, D. C., K. M. Kockelman, and N. Jiang. Animal-Vehicle Collisions in Texas: How to Protect Travelers and Animals on Roadways. *Accident Analysis & Prevention*, Vol. 131, 2019, pp. 157–170. https://doi.org/10.1016/j.aap.2019.05.030.
14. Ashraf, M. T., and K. Dey. Application of Bayesian Space-Time Interaction Models for Deer-Vehicle Crash Hotspot Identification. *Accident Analysis & Prevention*, Vol. 171, 2022, p. 106646. https://doi.org/10.1016/j.aap.2022.106646.
15. Moghaddam, K., V. Balali, P. Singh, and M. Khalilikhah. Evaluation of Multi-Class Multi-Label Machine Learning Methods to Identify the Contributing Factors to the Severity of Animal-Vehicle Collisions. *International Journal for Traffic and Transport Engineering*, Vol. 11, No. 3, 2021, pp. 341–358. https://doi.org/10.7708/ijtte2021.11(3).01.
16. Al-Bdairi, N. S. S., A. Behnood, and S. Hernandez. 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*, Vol. 26, 2020, p. 100120. https://doi.org/10.1016/j.amar.2020.100120.
17. Ahmed, S. S., J. Cohen, and P. C. Anastasopoulos. A Correlated Random Parameters with Heterogeneity in Means Approach of Deer-Vehicle Collisions and Resulting Injury-Severities. *Analytic Methods in Accident Research*, Vol. 30, 2021, p. 100160. https://doi.org/10.1016/j.amar.2021.100160.
18. Gharraie, I., and E. Sacchi. Severity Analysis of Wildlife–Vehicle Crashes Using Generalized Structural Equation Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 2021. 2675: 53–64.
19. Abra, F. D., B. M. Granziera, M. P. Huijser, K. M. P. M. B. Ferraz, C. M. Haddad, and R. M. Paolino. Pay or Prevent? Human Safety, Costs to Society and Legal Perspectives on Animal-Vehicle Collisions in São Paulo State, Brazil. *PLoS One*, Vol. 14, No. 4, 2019, p. e0215152. https://doi.org/10.1371/journal.pone.0215152.
20. Cherry, C. C., S. Dietz, E. Sauber-Schatz, S. Russell, J. Proctor, and D. Buttke. Characteristics of Animal-Related Motor Vehicle Crashes in Select National Park Service Units—United States, 1990–2013. *Traffic Injury Prevention*, Vol. 20, No. 1, 2019, pp. 58–63. https://doi.org/10.1080/15389588.2018.1508835.
21. Backs, J. A. J., J. A. Nychka, and C. C. St. Clair. Low Audibility of Trains May Contribute to Increased Collisions with Wildlife. *Transportation Research Interdisciplinary Perspectives*, Vol. 13, 2022, p. 100516. https://doi.org/10.1016/j.trip.2021.100516.
22. Conway, S., R. V. Rea, G. Hesse, C. MacPhail, A. von der Gonna, and D. Spooner. Exploratory Analysis of Physical and Emotional Impacts and Use of Healthcare Services Following Moose and Deer Vehicle Collisions in North-Central British Columbia. *Journal of Transport & Health*, Vol. 24, 2022, p. 101333. https://doi.org/10.1016/j.jth.2022.101333.
23. Llagostera, P., C. Comas, and N. López. Modeling Road Traffic Safety Based on Point Patterns of Wildlife-Vehicle Collisions. *Science of the Total Environment*, Vol. 846, 2022, p. 157237. https://doi.org/10.1016/j.scitotenv.2022.157237.
24. Rea, R. V., C. J. Johnson, D. A. Aitken, K. N. Child, and G. Hesse. Dash Cam Videos on YouTube™ Offer Insights into Factors Related to Moose-Vehicle Collisions. *Accident Analysis & Prevention*, Vol. 118, 2018, pp. 207–213. https://doi.org/10.1016/j.aap.2018.02.020.
25. Brieger, F., J.-L. Kämmerle, R. Hagen, and R. Suchant. Behavioural Reactions to Oncoming Vehicles as a Crucial Aspect of Wildlife-Vehicle Collision Risk in Three Common Wildlife Species. *Accident Analysis & Prevention*, Vol. 168, 2022, p. 106564. https://doi.org/10.1016/j.aap.2021.106564.
26. LADOTD. Crash1: LADOTD Highway Crash List. Louisiana Department of Transportation and Development. http://www8.dotd.la.gov/crash1/. Accessed March 5, 2023.
27. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
28. Das, S., and X. Sun. Factor Association with Multiple Correspondence Analysis in Vehicle–Pedestrian Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2519: 95–103.
29. Das, S., A. Dutta, M. Jalayer, A. Bibeka, and L. Wu. 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*, Vol. 7, No. 2, 2018, pp. 114–123. https://doi.org/10.1016/j.ijtst.2018.02.001.
30. Song, Y., H. Zhou, Q. Chang, and M. Jalayer. Multiple Correspondence Analysis of Wrong-Way Driving Fatal Crashes on Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2021. 2675: 1312–1323.
31. Das, S., and X. Sun. Association Knowledge for Fatal Run-Off-Road Crashes by Multiple Correspondence Analysis. *IATSS Research*, Vol. 39, No. 2, 2016, pp. 146–155. https://doi.org/10.1016/j.iatssr.2015.07.001.
32. Husson, F., S. Lê, and J. Pagès. *Exploratory Multivariate Analysis by Example Using R*. Chapman and Hall/CRC Press, Boca Raton, FL.
33. Pande, A., and M. Abdel-Aty. Market Basket Analysis of Crash Data from Large Jurisdictions and Its Potential as a Decision Support Tool. *Safety Science*, Vol. 47, No. 1, 2009, pp. 145–154. https://doi.org/10.1016/j.ssci.2007.12.001.

34. Weng, J., J. Z. Zhu, X. Yan, and Z. Liu. Investigation of Work Zone Crash Casualty Patterns Using Association Rules. *Accident Analysis & Prevention*, Vol. 92, 2016, pp. 43–52. https://doi.org/10.1016/j.aap.2016.03.017.
35. Rahman, M. A., X. Sun, S. Das, and S. Khanal. 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*, Vol. 10, No. 2, 2021, pp. 167–183. https://doi.org/10.1016/j.ijtst.2020.12.003.
36. Das, S., X. Sun, S. Goel, M. Sun, A. Rahman, and A. Dutta. Flooding Related Traffic Crashes: Findings from Association Rules. *Journal of Transportation Safety & Security*, Vol. 14, No. 1, 2022, pp. 111–129. https://doi.org/10.1080/19439962.2020.1734130.
37. Hossain, A., X. Sun, R. Thapa, and J. Codjoe. Applying Association Rules Mining to Investigate Pedestrian Fatal Injury Crash Patterns Under Different Lighting Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 2022. 2676: 659–672.
38. Montella, A., M. Aria, A. D'Ambrosio, and F. Mauriello. Analysis of Powered Two-Wheeler Crashes in Italy by Classification Trees and Rules Discovery. *Accident Analysis & Prevention*, Vol. 49, 2012, pp. 58–72. https://doi.org/10.1016/j.aap.2011.04.025.
39. Rahman, M. A., S. Das, and X. Sun. Single-Vehicle Run-Off Road Crashes Because of Cellphone Distraction: Finding Patterns with Rule Mining. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 1261–1277.
40. Rahman, M. A., X. Sun, M. Sun, and D. Shan. Investigating Characteristics of Cellphone Distraction with Significant Tests and Association Rule Mining. *IATSS Research*, Vol. 45, No. 2, 2020, pp. 198-209. https://doi.org/10.1016/j.iatssr.2020.09.001.
41. Maimon, O., and L. Rokach, eds. *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag, New York, NY, 2005.
42. Agrawal, R., and R. Srikant. Fast Algorithms for Mining Association Rules. *Proc., 20th VLDB Conference*, Santiago de Chile, Chile, 1994.
43. Hossain, M. A., S. Das, and X. Sun. Single-Vehicle Run-Off Road Crashes Because of Cellphone Distraction: Finding Patterns with Rule Mining. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 1261–1277.
44. Putzu, N., D. Bonetto, V. Civallero, S. Fenoglio, P. G. Meneguz, N. Preacco, and P. Tizzani. Temporal Patterns of Ungulate-Vehicle Collisions in a Subalpine Italian Region. *Italian Journal of Zoology*, Vol. 81, No. 3, 2014, pp. 463–470. https://doi.org/10.1080/11250003.2014.945974.
45. Responsibility.org. State Facts: Louisiana Drunk Driving Statistics. Foundation for Advancing Alcohol Responsibility. https://www.responsibility.org/alcohol-statistics/state-map/state/Louisiana. Accessed February 25, 2023.
46. Rahman, M. A., S. Das, and X. Sun. Understanding the Drowsy Driving Crash Patterns from Correspondence Regression Analysis. *Journal of Safety Research*, Vol. 84, 2023, pp. 167–181. https://doi.org/10.1016/j.jsr.2022.10.017.
47. Huijser, M. P., E. R. Fairbank, and F. D. Abra. *The Reliability and Effectiveness of a Radar-Based Animal Detection System*. Center for Environmentally Sustainable Transportation in Cold Climates, Fairbanks, AK.
48. Hardy, A., S. Lee, and A. F. Al-Kaisy. Effectiveness of Animal Advisory Messages on Dynamic Message Signs as a Speed Reduction Tool: Case Study in Rural Montana. *Transportation Research Record: Journal of the Transportation Research Board*, 2006. 1973: 64–72.