Journal of Substance Use

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ijsu20

Pattern recognition from cyclist under influence (CUI) crash events: application of block cluster analysis

Subasish Das, Kakan Dey & Md Tawhidur Rahman

To cite this article: Subasish Das, Kakan Dey & Md Tawhidur Rahman (2021): Pattern recognition from cyclist under influence (CUI) crash events: application of block cluster analysis, Journal of Substance Use, DOI: 10.1080/14659891.2021.1967483

To link to this article: https://doi.org/10.1080/14659891.2021.1967483

Published online: 19 Aug 2021.

JOURNAL OF SUBSTANCE USE
https://doi.org/10.1080/14659891.2021.1967483

Taylor & Francis
Taylor & Francis Group

Pattern recognition from cyclist under influence (CUI) crash events: application of block cluster analysis

Subasish Das [a], Kakan Dey [b], and Md Tawhidur Rahman [b]

[a]Traffic Operations and Roadway Safety Division, Texas A&M Transportation Institute, Bryan, Texas, USA; [b]Department of Civil and Environmental Engineering, West Virginia University, Morgantown, West Virginia, USA

ABSTRACT

Background: Alcohol impairment in traffic crashes is a critical safety concern. Alcohol impairment in non-motorist crashes has been rising in recent years. However, there is not much research focused on cyclist under influence (CUI).

Method: This study applied block cluster analysis to Louisiana traffic crash data from 2010 to 2016 to identify the key contributing attributes and association patterns of CUI crashes.

Results: The findings identified eight column clusters: hit and run crashes during cloudy conditions, impaired cyclist crashes in open country locations, younger impaired cyclist crashes at lighted intersections, elderly cyclist crashes during inclement weather, intersection crashes on low-speed roadways, segment-related crashes on undivided two-way roadways, fatal cyclist crashes at dark with no lighting, and collision with a vehicle on business locality.

Conclusions: The findings of this study can be beneficial in the synchronization of regional and local behavioral safety efforts to lessen the occurrence and injury level of CUI crashes.

ARTICLE HISTORY
Received 15 February 2021
Accepted 6 August 2021

KEYWORDS
Cyclist crashes; cyclist impairment; pattern recognition; block cluster analysis

Introduction

Cycling has become one of the popular transportation modes for short duration trips in urban areas. As cyclists are increasing in numbers, there is a possibility of an increase in cyclist involved crashes. Alcohol is one of the major risk factors for traffic crashes because it can impair judgment and cognitive function (Stübig et al., 2012). Many researchers have focused on the safety-related issues of cyclist. However, only a limited number of studies focused on alcohol impairment among pedestrians and cyclists in the United States. There were 857 cyclist deaths in 2018 (2.3% of all traffic fatalities in 2017). From 2009 to 2018, the average age of cyclists killed in motor vehicle crashes has increased from 41 to 47. Alcohol involvement (BAC of .01 + g/dL) was present in 37% of all cyclist fatalities in 2018. An estimated 20% of fatal cyclist crashes had a cyclist with a BAC of .08 g/dL or higher (NHTSA, 2020a, 2020b). These statistics indicate a need for an in-depth study to identify the patterns and associations between key contributing factors from cyclists under the influence (CUI) crashes.

This study aims to answer three key research questions (**RQ**): 1) **RQ1**: What are the key contributing factors of CUI crashes? 2) **RQ2**: Do the explanatory variables different by crash injury types in these crashes? 3) **RQ3**: What are the key clusters of variable attributes in CUI crashes? This study acquired seven years (2010–2016) of traffic crash data from Louisiana. After performing some preliminary investigations, this study applied a block clustering approach to identify the association patterns of the key contributing factors.

Literature review

It is important to understand the key factors of CUI crashes in order to improve roadway safety. A comprehensive understanding of key attributes and patterns of these crashes is critical to the development of effective safety strategies and countermeasures. There are several studies focused on CUI crashes, and they are summarized below.

Using five years (1987–1991) of Fatal Accident Reporting System (FARS) data, Li and Baker (1994) analyzed blood alcohol concentrations (BACs) among fatally injured cyclists. Cyclists aged 25 to 34 who died in nighttime crashes had a significantly increased likelihood of being BAC positive and legally intoxicated. Additionally, 14% of cyclists in the younger group (ages 15 to 19 years) had positive BACs. Based on trauma registry data (1990–1997), Li et al. (2000) assembled a historical cohort of 120 Maryland residents (with bicycle crash history) aged 18 years or older. Cyclists with positive BACs were significantly more likely to have a record of license suspension/revocation (52% vs. 14%, p-value < 0.01) and to have driving under influence (DUI) convictions (30% vs. 3%, p-value < 0.01) when compared to those with negative BACs

Eichelberger et al. (2018) investigated the prevalence, trends, and characteristics of alcohol-impaired fatal pedestrian and bicycle crashes. DiMaggio et al. (2016) analyzed the spatial risk of alcohol-related pedestrian/cyclist injury in New York City at the U.S. census tract level for a recent 10-year period using a Bayesian hierarchical spatial regression model with Integrated Nested Laplace approximations. The findings show that the presence of at least one alcohol outlet in a census tract increased the risk of a pedestrian or cyclist being struck by a car

CONTACT Subasish Das s-das@tti.tamu.edu Texas A&M Transportation Institute, 1111 RELLIS Parkway, Bryan, TX 77807, USA

by 47%. Kwigizile et al. (2014) examined six years of crash data in Jacksonville, Florida to determine the impact of alcohol and drug use on cyclist injuries in traffic crashes. The study found that significant variables of the injury outcome for crashes involving alcohol or drug use are not necessarily significant for crashes not involving alcohol or drug use.

The literature review shows that a comprehensive study of CUI crashes is needed. The present study shows how innovative dimension reduction methods can gain valuable insights into CUI crashes.

Methodology

Data collection

This study used crash data from the Louisiana Department of Transportation and Development (LADOTD), spanning from 2010 to 2016. For the database, the information was divided into three levels: crash level, vehicle level, and road inventory level. Regarding the vehicle level, this study only considers crashes with bicycle as the vehicle type. After filtering all bicycle crashes, a variable code (Est_Alcohol) is used to label the CUI crashes. The "Est_Alcohol" includes two categories: 0 indicates an impaired crash, and 1 indicates a non-impaired crash. This variable has been generated from a precise mathematical model capable of classifying whether or not a driver involved in a crash was under the influence of alcohol. Table 1 lists the frequencies of bicycle crashes, and CUI crashes. Fatalities related to bicycle and CUI crashes are also listed in Table 1. From 2010 to 2016, the number of bicycle crashes increased by 43% (for traffic fatalities, this increase is 133%). There was a sudden rise of CUI crashes and fatalities in 2015.

Table 1. Crashes involving bicycles in Louisiana (2010–2016).

Year	Bicyclist Crashes	CUI Crashes	Bicyclist Fatalities	CUI Fatalities
2010	631	38	9	9
2011	847	42	16	5
2012	912	36	24	6
2013	927	38	13	4
2014	882	41	12	5
2015	876	56	33	12
2016	906	37	21	4

Block cluster analysis

Categorical data analysis, such as cluster analysis and correspondence analysis, have been becoming popular in transportation safety analysis due to the complexity of crash data and the abundance of categorical information in crash data (Das et al., 2018, 2020; Jalayer et al., 2018). This section provided a brief overview of the concept of block clustering, without providing all theoretical details, based on several studies (Das et al., 2020; Govaert & Nadif, 2008, 2013; Madeira & Oliveira, 2004). Interested readers can consult these papers (Das et al., 2020; Govaert & Nadif, 2008, 2013) to get a more comprehensive understanding of this concept.

Block clustering, a clustering technique, simultaneously examines the two sets, observations, and variables and organizes the data into homogeneous blocks. Suppose Xdenotes an $n \times d$ data matrix defined by $X = (xij)$; i & j $\in$ J}, where I is a set of n objects (rows, observations, crashes) and J is a set of $d$variables (columns, variables, attributes). The main goal of this method is to generate permutations or rearrangements of observations and characteristics to construct a correspondence structure on $I \times J$. The one significant advantage regarding block clustering is the transformation of the initial data matrix X into a smaller and less complex data matrix with the same structure. Figure 1 shows an example of block clustering.

Results

Exploratory data analysis

The final dataset of this study contains 288 crash level information. To compute the count and percentages of variable categories by groups (i.e., crash injury type in this case), it is important to test whether the distribution

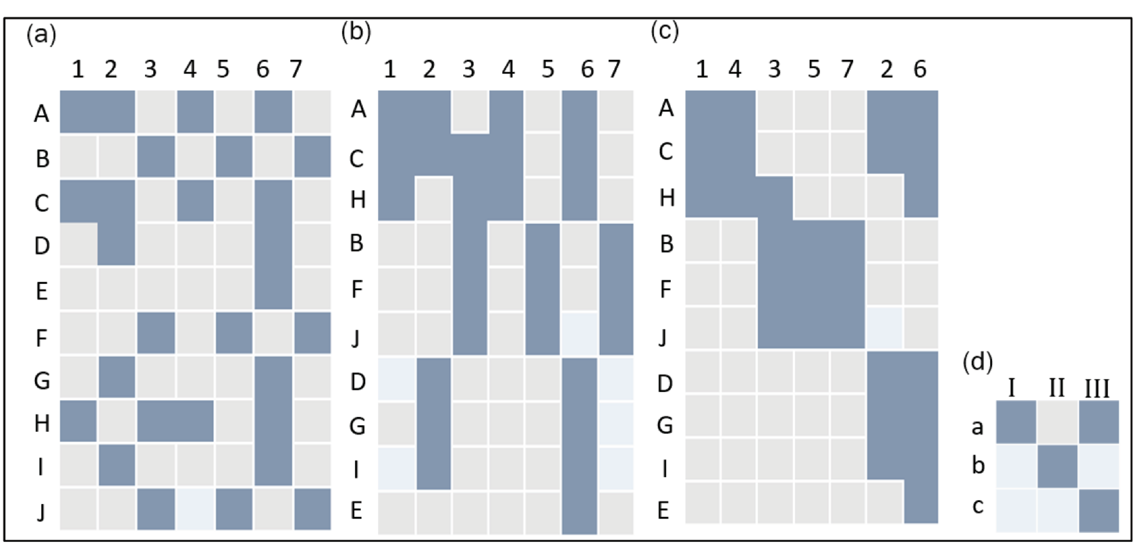

Figure 1. Block clustering, showing (a) binary data set, (b) data reorganized by a partition on I, (c) data reorganized by partitions on I and J simultaneously, and (d) summary binary data.

of the variables differs between groups. Statistical analyses were performed using R (version 3.6.0) using the package "compareGroups" (Salvador, 2020) for descriptive tables. This study defined statistical significance as p-value < 0.05. Five crash injury levels were used in this study: 1) fatal (K), 2) severe (A), 3) moderate (B), 4) complaint (C), and 5) non injury (O). After performing a co-relation analysis and variable importance analysis, ten variables were selected for analysis. As all of the variables of interest in this group are categorical, a Chi-square test was performed. As this study design contains more than two groups, there is a need for performing overall association assessment as well as pairwise comparisons. This study conducted pairwise tests and displayed p-values (as listed in Table 2). From Table 2, the number of involved vehicles, presence of intersection, and roadway type were found to be the only three covariates that significantly differ by crash injury types (at 90% confidence boundary). The results indicate that CUI requires an innovative data analysis process instead of conducting a multinomial injury severity analysis.

Results of block cluster analysis

n order to determine the contributing factors of CUI crashes, these crashes can be compared with non-CUI crashes. The cyclist-involved datasets have a larger number of non-CUI crashes than CUI crashes. However, under sampling and over sampling using disproportionate datasets usually are not insightful. The current analysis is limited to the CUI crash dataset only.

The initial step for block clustering is to identify the optimal number of clusters for columns and rows. This study used an integrated completed likelihood (ICL) model to maximize the complete data likelihood. At first, a wide range of combinations (1 to 20 for rows and 1 to 20

Table 2. Descriptive statistics by crash severity types.

Variable Category	K (N = 52)	A (N = 19)	B (N = 86)	C (N = 94)	O (N = 37)	p-value
DR_AGE (Driver Age)						n.s.*
< 20 yrs	5 (9.62%)	1 (5.26%)	1 (1.16%)	0 (0.00%)	0 (0.00%)	
20–39 yrs	15 (28.8%)	2 (10.5%)	15 (17.4%)	22 (23.4%)	12 (32.4%)	
40–59 yrs	26 (50.0%)	13 (68.4%)	60 (69.8%)	63 (67.0%)	23 (62.2%)	
> 59 yrs	6 (11.5%)	3 (15.8%)	10 (11.6%)	9 (9.57%)	2 (5.41%)	
HIT_AND_RUN (Hit and Run)						n.s.
No	38 (73.1%)	16 (84.2%)	70 (81.4%)	76 (80.9%)	30 (81.1%)	
Yes	14 (26.9%)	3 (15.8%)	16 (18.6%)	18 (19.1%)	7 (18.9%)	
NUM_VEH (Vehicles Involved)						< 0.01
Single Bike	0 (0.00%)	0 (0.00%)	1 (1.16%)	2 (2.13%)	3 (8.11%)	
Single Veh	46 (88.5%)	19 (100%)	81 (94.2%)	90 (95.7%)	33 (89.2%)	
Multi Veh	6 (11.5%)	0 (0.00%)	4 (4.65%)	2 (2.13%)	1 (2.70%)	
INTERSECTION (Intersection)						< 0.01
No	36 (69.2%)	16 (84.2%)	49 (57.0%)	59 (62.8%)	13 (35.1%)	
Yes	16 (30.8%)	3 (15.8%)	37 (43.0%)	35 (37.2%)	24 (64.9%)	
LOC_TYPE_CD (Locality Type)						n.s.
Residential	16 (30.8%)	10 (52.6%)	25 (29.1%)	27 (28.7%)	8 (21.6%)	
Business	32 (61.5%)	9 (47.4%)	57 (66.3%)	64 (68.1%)	28 (75.7%)	
Open Country	3 (5.77%)	0 (0.00%)	2 (2.33%)	2 (2.13%)	1 (2.70%)	
Other	1 (1.92%)	0 (0.00%)	2 (2.33%)	1 (1.06%)	0 (0.00%)	
HWY_TYPE_CD (Highway Type)						n.s.
City Street/Parish Road	15 (28.8%)	8 (42.1%)	43 (50.0%)	49 (52.1%)	20 (54.1%)	
State Hwy	25 (48.1%)	6 (31.6%)	29 (33.7%)	35 (37.2%)	10 (27.0%)	
U.S. Hwy	9 (17.3%)	5 (26.3%)	12 (14.0%)	10 (10.6%)	5 (13.5%)	
Not reported	3 (5.77%)	0 (0.00%)	2 (2.33%)	0 (0.00%)	2 (5.41%)	
ROAD_TYPE_CD (Road Type)						< 0.01
Two-Way Road with No Physical Separation	28 (53.8%)	15 (78.9%)	53 (61.6%)	55 (58.5%)	17 (45.9%)	
Two-Way Road with A Physical Barrier	2 (3.85%)	0 (0.00%)	0 (0.00%)	2 (2.13%)	0 (0.00%)	
Two-Way Road with A Physical Separation	16 (30.8%)	2 (10.5%)	22 (25.6%)	20 (21.3%)	10 (27.0%)	
One-Way Road	6 (11.5%)	2 (10.5%)	10 (11.6%)	15 (16.0%)	9 (24.3%)	
Other	0 (0.00%)	0 (0.00%)	1 (1.16%)	2 (2.13%)	1 (2.70%)	
PSL (Posted Speed Limit)						n.s.
< 40 mph	20 (38.5%)	9 (47.4%)	54 (62.8%)	60 (63.8%)	33 (89.2%)	
40–50 mph	19 (36.5%)	7 (36.8%)	23 (26.7%)	24 (25.5%)	3 (8.11%)	
> 50 mph	13 (25.0%)	3 (15.8%)	9 (10.5%)	10 (10.6%)	1 (2.70%)	
WEATHER_CD (Weather)						n.s.
Clear	33 (63.5%)	12 (63.2%)	77 (89.5%)	75 (79.8%)	28 (75.7%)	
Cloudy	14 (26.9%)	5 (26.3%)	5 (5.81%)	16 (17.0%)	6 (16.2%)	
Inclement	5 (9.62%)	2 (10.5%)	4 (4.65%)	3 (3.19%)	3 (8.11%)	
DAY_OF_WK (Day of Week)						n.s.
FSS	24 (46.2%)	7 (36.8%)	41 (47.7%)	39 (41.5%)	18 (48.6%)	
MTWT	28 (53.8%)	12 (63.2%)	45 (52.3%)	55 (58.5%)	19 (51.4%)	

** n.s. = not significant*

for columns) were tried to find the optimum number of blocks by examining ICL and pseudolikelihood values for each run. The optimum number of blocks was found to be 40, with 5 rows and 8 columns.

Discussions

Table 3 lists the attributes for each of the column clusters. Discussion on each of the clusters is provided in this section.

Cluster 1

The major variable categories in this cluster are hit and run crashes, U. S. highways, cloudy weather, driver condition as the most harmful event, posted speed limits of 40–50 mph, sideswipe crashes, and roadways with a white dashed line. This particular cluster shows the attribute patterns that are associated with hit and run bicycle crashes. The attribute group indicates that hit and run crashes (sideswiped crashes in cloudy condition) mostly occur due to the driver condition. The most harmful event for these crashes is unknown. These crashes are generally associated with moderate speed limits (40–50 mph) on U.S. highways with white dashed lines as the traffic control tool.

Cluster 2

The major variable categories of this cluster are dawn/dusk as the lighting condition, open country as the locality, avoiding other objects/vehicle failure/vehicle out of control as the most harmful event, right-turn collisions, two-way roads with a physical barrier, single-bike crashes, and multiple-vehicle crashes. The cyclists in this group are mostly younger. The clusters indicate that right-turn crashes during dawn or dusk on open country roadways are associated with young, impaired cyclists. Li and Baker (1994) found that it was in the age range of 25–34 that the largest number of crashes involved cyclists tested positive for alcohol.

Cluster 3

This cluster includes variables such as drivers aged 20 to 39 years, moderate injury or complaint, lighted intersection at night, low-speed roadways, city streets or Parish roads, roadways with no traffic control, right-angle crashes at residential localities, normal condition or driver violation as the most harmful event, and two-way roadways with physical separation. This cluster indicates that young, impaired cyclist-related right-angle crashes at lighted intersections are associated with moderate injuries or complaints. Within urban environment (city streets and Parish roads), alcohol consumption by cyclists is associated with more moderate and complaint injuries.

Cluster 4

The variable categories in this cluster are weekdays, city streets or Parish roads, intersections, lower speed limits, daylight, and right-angle crashes. This cluster shows the CUI crashes are

Table 3. Column clusters.

Variable Category	Cluster	Variable Category	Cluster
HIT_AND_RUN_Yes	1	DAY_OF_WK_MTWT	4
HWY_TYPE_CD_U.S. Hwy	1	HWY_TYPE_CD_City Street/Parish Road	4
M_HARM_EV_CD_Due to Driver Condition	1	INTERSECTION_Yes	4
M_HARM_EV_CD_Unknown	1	LIGHTING_CD_Daylight	4
MAN_COLL_CD_Sideswipe	1	MAN_COLL_CD_Right Angle	4
PSL_40-50 mph	1	PSL_< 40 mph	4
TRAFF_CNTL_CD_White Dashed Line	1	DR_AGE_> 59 yrs	5
WEATHER_CD_Cloudy	1	DR_INJ_CD_A	5
DR_AGE_< 20 yrs	2	DR_INJ_CD_O	5
HWY_TYPE_CD_Not reported	2	LIGHTING_CD_Dark – Street Light at Intersection Only	5
LIGHTING_CD_Dusk/Dawn	2	MAN_COLL_CD_Head-On	5
LIGHTING_CD_Not reported	2	MAN_COLL_CD_Left Turn	5
LOC_TYPE_CD_Open Country	2	MAN_COLL_CD_Other	5
LOC_TYPE_CD_Other	2	MAN_COLL_CD_Single Bike	5
M_HARM_EV_CD_Due to Vehicle Condition (Failure)	2	ROAD_TYPE_CD_One-Way Road	5
M_HARM_EV_CD_NA	2	TRAFF_CNTL_CD_Not reported	5
M_HARM_EV_CD_Other	2	TRAFF_CNTL_CD_Other	5
M_HARM_EV_CD_To Avoid Other Object	2	TRAFF_CNTL_CD_Signal	5
M_HARM_EV_CD_To Avoid Other Vehicle	2	TRAFF_CNTL_CD_Stop Sign	5
M_HARM_EV_CD_Vehicle Out of Control, Not Passing	2	WEATHER_CD_Inclement	5
MAN_COLL_CD_Right Turn	2	DAY_OF_WK_FSS	6
NUM_VEH_Multi Veh	2	HWY_TYPE_CD_State Hwy	6
NUM_VEH_Single Bike	2	INTERSECTION_No	6
ROAD_TYPE_CD_Other	2	ROAD_TYPE_CD_Two-Way Road with No Physical Separation	6
ROAD_TYPE_CD_Two-Way Road with a Physical Barrier	2	DR_INJ_CD_K	7
DR_AGE_20-39 yrs	3	LIGHTING_CD_Dark – No Street Lights	7
DR_INJ_CD_B	3	MAN_COLL_CD_Rear End	7
DR_INJ_CD_C	3	PSL_> 50 mph	7
LIGHTING_CD_Dark – Continuous Street Light	3	TRAFF_CNTL_CD_Yellow Dashed Line	7
LOC_TYPE_CD_Residential	3	TRAFF_CNTL_CD_Yellow No Passing Line	7
M_HARM_EV_CD_Due to Driver Violation	3	DR_AGE_40-59 yrs	8
M_HARM_EV_CD_Normal Movement	3	HIT_AND_RUN_No	8
ROAD_TYPE_CD_Two-Way Road with A Physical Separation	3	LOC_TYPE_CD_Business	8
TRAFF_CNTL_CD_No Control	3	NUM_VEH_Single Veh	8
		WEATHER_CD_Clear	8

clustered based on the conventional urban intersection locations with a high propensity for right-angle crashes. The results are in line with DiMaggio et al. (2016) study.

Cluster 5

The variable categories in this cluster are elderly cyclists, severe injury, no injury, darkness with streetlights at intersections only, single-bike crashes, left-turn or head-on collisions, inclement weather, one-way roads, and traffic signals or stop signs as the traffic control devices. These results are in line with Das et al. (2019) study. A host of risk-taking behaviors due to alcohol consumption can lead cyclists to more risk-taking behavior such as riding at night and crossing at intersections with no lighting. Juhra et al. (2012) reported that cyclists often disregard traffic laws while crossing, which was one of the main causes of bicycle collisions with motor vehicles.

Cluster 6

The major variables categories in this cluster are weekends, state highways, road segments, and two-way roads with no physical separation. Two-way roads with no physical separation is a broad category, so this cluster limited these roadways to state highways and segment related. Adding bicycle lanes and widening shoulders are the two common treatments to reduce these crashes.

Cluster 7

The major variables categories in this cluster are fatal cyclist crashes, darkness with no lighting, rear-end collisions, high-speed roadways, and yellow dashed lines or yellow with no passing as the traffic control device. This cluster clearly shows that fatal cyclist crashes are associated with dark with no lighting locations. Many studies identified "dark no lighting" as a key contributing factor in non-motorist crashes (DiMaggio et al., 2016; Das & Sun, 2015; Das et al., 2019). Installation of lighting at crash-prone locations could be a suitable countermeasure.

Cluster 8

The major variables categories in this cluster are drivers aged 40 to 59 years old, impaired cyclists, single-vehicle collisions with a cyclist, clear weather, and business locality. This cluster is not associated with hit and run crashes. This cluster represents a general trend of conventional bicycle crash scenario at a business location with no specific traits.

This study identified eight clusters of factors that represent the association patterns of CUI crashes. Findings of this study can be used by different transportation agencies and appropriate policy and enforcement related decisions can be made to reduce these crashes.

Limitations

There are a number of limitations to this study that should be acknowledged. First, this study incorporated a limited number of factors that are associated with CUI crashes. Some variables, such as roadway traffic condition, driver action, and roadway visibility, had a high number of missing values and, therefore, were not included. Another limitation of the current study is that it does not utilize comprehensive crash typing data for fatal crashes. A third limitation is that the analysis is solely based on police-reported crashes. Limitation of this study can be mitigated in the future studies.

Conclusions

Cycling has become increasingly popular in the U.S. The increased number of cyclists has also led to an increased number of bicycle crashes, especially in dense urban locations. Regular cyclists and bikeshare users often use bicycles to return home from parties or other social gatherings where alcohol serving has been allowed. The key reason behind CUI is the reluctance of the cyclist in driving a car after the consumption of alcohol. Whilst CUI does pose less threat to other road users, the rider poses a great threat to himself. This study examined the applicability of block clustering methods by analyzing seven years of Louisiana CUI crash data. Instead of conducting crash severity analysis, this study used an innovative method that can produce both column and row-based clusters. The results show that this method can provide a more comprehensive interpretation of the explanatory variables' effects and interactions by showing different magnitudes among clusters. The findings identified eight column clusters: hit and run crashes during cloudy conditions, impaired cyclist crashes in open country locations, younger impaired cyclist crashes at lighted intersections, elderly cyclist crashes during inclement weather, intersection crashes on low-speed roadways, segment related crashes on undivided two-way roadways, fatal cyclist crashes at dark with no lighting, and collision with a vehicle on business locality. To prevent CUI crashes, alternative ways of transportation (such as public transport or the designated driver concept or using ride sharing services such as Uber and Lyft after alcohol consumption) must be promoted as well. Proactive regulations such as utilization of a protective helmet can help in decreasing injury levels of the cyclists.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Subasish Das http://orcid.org/0000-0002-1671-2753
Kakan Dey http://orcid.org/0000-0002-5875-6180

References

Das, S., Avelar, R., Dixon, K., & Sun, X. (2018). Investigation on the wrong way driving crash patterns using multiple correspondence analysis. *Accident Analysis & Prevention*, *111*, 43–55. https://doi.org/10.1016/j.aap.2017.11.016

Das, S., Jha, K., Fitzpatrick, K., Brewer, M., & Shimu, T. H. (2019). Pattern identification from older bicyclist fatal crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *2673* (6), 638–649. https://doi.org/10.1177/0361198119841863

Das, S., & Sun, X. (2015). Factor association with multiple correspondence analysis in vehicle–pedestrian crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *2519*(1), 95–103. https://doi.org/10.3141/2519-11

Das, S., Tran, L. N., & Theel, M. (2020). Understanding patterns in Marijuana impaired traffic crashes. *Journal of Substance Use*, *26*(1), 21–29. https://doi.org/10.1080/14659891.2020.1760381

DiMaggio, C., Mooney, S., Frangos, S., & Wall, S. (2016). Spatial analysis of the association of alcohol outlets and alcohol-related pedestrian/bicyclist injuries in New York city. *Injury Epidemiology*, *3*(1), 10. https://doi.org/10.1186/s40621-016-0076-5

Eichelberger, A. H., McCartt, A. T., & Cicchino, J. B. (2018). Fatally injured pedestrians and cyclists in the United States with high blood alcohol concentrations. *Journal of Safety Research*, *65*, 1–9. https://doi.org/10.1016/j.jsr.2018.02.004

Govaert, G., & Nadif, M. (2008). Block clustering with bernoulli mixture models: Comparison of different approaches. *Computational Statistics & Data Analysis*, *52*(6), 3233–3245. https://doi.org/10.1016/j.csda.2007.09.007

Govaert, G., & Nadif, M. (2013). *Co-clustering: Models, algorithms and applications* (1st ed.). Wiley-IEEE Press.

Jalayer, M., Zhou, H., & Das, S. (2018). *Exploratory analysis of run-off-road crash patterns. Data analytics for smart cities*. CRC Press.

Juhra, C., Wieskötter, B., Chu, K., Trost, L., Weiss, U., Messerschmidt, M., Malczyk, A., Heckwolf, M., & Raschke, M. (2012). Bicycle accidents e do we only see the tip of the iceberg? A prospective multicentre study in a large German city combining medical and police data. *Injury*, *43*(12), 2026–2034. https://doi.org/10.1016/j.injury.2011.10.016

Kwigizile, V., Sando, T., & Chimba, D. (2014). Understanding the role of alcohol and drugs on bicyclist injury severity. *Advances in Transportation Studies*, *34*, 43–56. https://www.atsinternationaljournal.com/index.php/2014-issues/xxxiv-november-2014/562-understanding-the-role-of-alcohol-and-drugs-on-bicyclist-injury-severity

Li, G., Shahpar, C., Soderstrom, C. A., Baker, S. P., & Elsevier. (2000). Alcohol use in relation to driving records among injured cyclists. *Accident Analysis & Prevention*, *32*(4), 583–587. https://doi.org/10.1016/S0001-4575(99)00089-5

Li, G. S., & Baker, P. (1994). Alcohol in fatally injured cyclists. *Accident Analysis & Prevention*, *26*(4), 543–548. https://doi.org/10.1016/0001-4575(94)90045-0

Madeira, S. C., & Oliveira, A. L. (2004). Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *1*(1), 24–45. https://doi.org/10.1109/TCBB.2004.2

NHTSA. 2020a. *Traffic safety facts: 2018 data (cyclists and other cyclists)*. Report No. DOT HS 812 884. Retrieved August 1, 2020a, from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812884

NHTSA. 2020b. *Traffic safety facts: 2017 data (cyclists and other cyclists)*. Report No. DOT HS 812 765. Retrieved August 1, 2020b, from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812765

Salvador, I. S. J. 2020. *R package 'compareGroups*. Retrieved August 1, 2020, from https://cran.r-project.org/web/packages/compareGroups/compareGroups.pdf

Stübig, T., Petri, M., Zeckey, C., Brand, S., Müller, C., Otte, D., Krettek, C., & Haasper, C. (2012). Alcohol intoxication in road traffic accidents leads to higher impact speed difference, higher ISS and MAIS, and higher preclinical mortality. *Alcohol*, *46*(7), 681–686. https://doi.org/10.1016/j.alcohol.2012.07.002