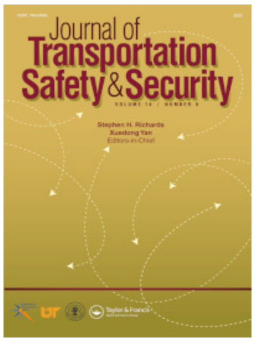


Journal of Transportation Safety & Security

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/utss20

Pattern recognition from injury severity types of frontage roadway crashes

Subasish Das, Reuben Tamakloe, Boniphace Kutela & Ahmed Hossain

To cite this article: Subasish Das, Reuben Tamakloe, Boniphace Kutela & Ahmed Hossain (2022): Pattern recognition from injury severity types of frontage roadway crashes, Journal of Transportation Safety & Security, DOI: 10.1080/19439962.2022.2123581

To link to this article: https://doi.org/10.1080/19439962.2022.2123581

Published online: 14 Sep 2022.

JOURNAL OF TRANSPORTATION SAFETY & SECURITY
https://doi.org/10.1080/19439962.2022.2123581

Pattern recognition from injury severity types of frontage roadway crashes

Subasish Das[a], Reuben Tamakloe[b], Boniphace Kutela[c], and Ahmed Hossain[d]

[a]Ingram School of Engineering, Texas State University, San Marcos, Texas, USA; [b]Department of Transportation Engineering, The University of Seoul, Seoul, South Korea; [c]Texas A&M Transportation Institute, Texas A&M University System, College Station, Texas, USA; [d]Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, Louisiana, USA

ABSTRACT

Frontage roads are the supporting roadways that are along freeways and fully controlled principal arterial roadway networks in the U.S. These roads are designed in a way to provide access between the freeways, principal arterials, and surrounding business entities. For Texas, these roadways are the leading design resolution for providing access along rural freeways and principal arterial roadways. These roadways are generally two-ways for rural and less developed urban areas and are mostly one-way for urban and city-centered roadways. Although frontage roadways possess major safety concerns, the safety performance of these roadways has not been well studied. This study collected six years of frontage road crash data from Texas to determine the patterns of associated factors by applying a dimension reduction method known as cluster correspondence analysis (CCA). The results revealed four clusters for each of the two datasets based on crash injury types. For fatal and injury crashes, the major clusters are distraction-related crashes at signalized intersections, segment-related crashes at dark unlighted conditions, yield signed intersection locations and segments with no TCDs, and intersection crashes on undivided roadways. For the no injury crash dataset, the key clusters are segment crashes in dark conditions and rain, crashes at signalized intersections with both drivers going straight, segment crashes with both drivers going straight with marked lanes or no TCDs, and intersection-related collisions on undivided roadways. Based on the evaluation results, suitable safety countermeasures and policy initiatives to reduce frontage road crash frequencies can be singled out.

KEYWORDS
frontage road; traffic crashes; crash severity; cluster correspondence analysis

Introduction

Frontage roads are the supporting roadways that run along freeways and fully controlled principal arterial corridors in the U.S. The major function of these roadways is to provide access to the higher functional class

CONTACT Subasish Das subasish@txstate.edu Ingram School of Engineering, Texas State University, 601 University Drive, San Marcos, TX 78666, USA

roadways and adjacent properties on both sides of the arterial corridors (Lord & Bonneson, 2007). Frontage roadways are widely popular in Texas, and both one and two-way frontage roadways are commonly seen. Rural and less developed locations have two-way frontage roads, and urban and city area locations typically have these roadways as one-way on both sides of the major functional class roadways. Numerous conflicting points, due to the abundance of driveways, intersections, and exit and entrance areas, make these roadways crash-prone (Machsus, Prayogo, Hayati, & Utanaka, 2017). However, the safety investigation of these roadways has not been adequately explored (Lord & Bonneson, 2007).

In traffic safety research, the majority of studies focused on identifying the effect of single attributes on the injury severities resulting from crashes. Nevertheless, combinations of attributes are most likely to collectively influence the severity levels of crashes. Given the critical nature of traffic flows on frontage roads, attributes such as poor visibility and driver distraction and maneuver could collectively contribute to a crash. Using parametric models would not be able to identify these groups of risk factors for the identification of proper countermeasures (Tamakloe, Das, Nimako Aidoo, & Park, 2022). Furthermore, the dataset used in this study has many categorical variables with a large set of attributes or categories. Traditionally, the analysis would involve performing correlation analysis and selecting variables that are not highly correlated to avoid the perfect multicollinearity (Daoud, 2017; Kutela, 2022). However, such approach would eliminate the potential of variables (with categories per variable), which could provide insightful patterns. To the best knowledge of the researchers, no study has investigated the group of factors jointly influencing the severity of injuries sustained in crashes on frontage roads. Identifying these clusters of attributes using cluster analysis would provide a better understanding of factors influencing the severity of crashes and will inform policy decisions geared at reducing the severity of crashes on frontage roads.

Against this backdrop, this study aims to mitigate this critical research gap by examining the critical clusters that describe the patterns among attributes influencing the injury severity of crashes occurring on frontage roads using an advanced unsupervised data mining algorithm. To achieve this goal, the study collected six years (2014–2019) of frontage road-related crash data from Texas in order to determine the association between key crash contributing factors. To determine the robust association patterns, this study employed a less explored dimension reduction method known as cluster correspondence analysis (CCA). Unlike the conventional statistical methods that heavily depend on prior assumptions that may influence the results of the analysis, CCA has no pre-assumptions and basically identifies the cluster of attributes leading to an event. Besides, unlike other clustering methods such as the simple cluster analysis, the dimension reduction method in CCA simultaneously assigns attributes to clusters using an optimal scaling process that maximize the between-group variation.

CCA has been employed in safety analysis to determine the group of factors that occur together in order to identify the risk factor patterns in a large dataset (Das, Avelar, Dixon, & Sun, 2018; Das, Chatterjee, et al., 2021). For a crash dataset with a large number of categorical variables, CCA methods, such as multiple correspondence analysis (MCA), are commonly used (Das et al., 2018). MCA clusters are mostly based on user understanding of the nearby located factors in a two-dimensional space. CCA is more suitable as the cluster determination is also performed by the algorithms. This study identified several key clusters, which provided latent insights on association patterns of frontage road-related crashes. Based on the evaluation results, suitable safety countermeasures and policy initiatives to reduce frontage road crash frequencies can be singled out.

Literature review

Road engineers have used frontage roads as a design solution for reducing traffic congestion on highways as they serve as a transition between the high-speed traffic on the highways and the low-speed traffic on local streets. Owing to the changing dynamics regarding land use near frontage roads and the traffic congestion associated with them, they have become common areas for crashes (Kockelman et al., 2021; Kusselson, 2013; Xu, Kockelman, & Wang, 2014). In particular, Bhat, Born, Sidharthan, and Bhat (2014) noted that crash rates on frontage roads in Texas are likely to be about 2.5 times higher relative to intersections not found on frontage roads. Zhang, Xie, and Li (2012) also identified that frontage roads, particularly those with diamond intersections, are more crash-prone compared to freeways and urban arterials. Even though these roads are prone to crashes and associated with severe injury outcomes, a review of the literature reveals that safety evaluation on frontage roadways is limited. This section provides a brief overview of the frontage road-related safety studies.

Research on frontage road safety has identified that these roads are particularly associated with high turning volumes and weaving traffic conflicts (conflicts at segments where merge intersections are directly accompanied by diverging intersections). According to the literature, critical frontage road segments include intersections, diamond interchanges, and weaving sections (Bhat et al., 2014; Mallipaddi & Anderson, 2020; Zhang et al., 2012) due to their potential for vehicle conflicts. A comprehensive study that explored crashes on freeway and frontage road weaving segments in Huntsville, Alabama between 2010 and 2017 identified that weaving segment-related crashes on frontage roads were mainly rear-end crashes and mostly occurred during clear weather conditions. Interestingly, females were found to have caused 60.4% of the weaving crashes on frontage roads, and crashes involving drivers aged 25 to 64 years, and most crashes occurring at night and on wet roads were severe. The authors also mentioned that weaving segment-related crashes on frontage roads were more severe when drivers failed to follow traffic rules (Mallipaddi & Anderson, 2020). Regarding crashes at intersections on frontage roads, Bhat et al. (2014) pointed out that intersections with a frontage road approach are particularly dangerous locations as motorists are less likely to reduce their speeds upon exiting the highways. The authors recommended driver education through campaigns to help reduce the number of crashes in these segments.

A variety of factors, including weather, roadway design, and environmental features, have been found to impact crash frequency and severity on frontage roads (Gattis, Hanning, & Duncan, 2008; Xu et al., 2014; Zhao, Goodman, Azimi, & Qi, 2018). According to a police report, the design of frontage roads in Texas, particularly at entry ramps, confuses truck drivers (Zhao et al., 2018). In addition, most truck drivers fail to see the yield signs at entry ramps when making left turn maneuvers due to the high speeds at these segments. According to the authors, crashes are likely to be severe head-on collisions. Xu et al. (2014) employed a wide variety of variables, such as crash frequencies, truck volumes, traffic intensity, land use, rainfall, local population density, income, job density, and education levels, to investigate the role of design, built environment, and weather on frontage road fatality counts. To address excess zero crash counts, a zero-inflated negative binomial (ZINB) model was developed. The results explained the association between crash frequencies with rainfall, local job densities, and population densities. In a similar study, Gattis et al. (2008) examined the safety and operational impact of frontage roads along Interstate 30 in central Arkansas and the transition from a two-way operation to a one-way operation. The factors considered in this analysis were crash frequencies, traffic volumes, operating speeds, land use patterns, travel times between different key points of interest, and sentiments of surrounding business owners.

Several other studies explored the use of crash modification factors (CMF's) in their analysis. Lord and Bonneson (2007) developed safety performance functions (SPF) and CMFs for rural frontage road segments. This study also examined one-way and two-way frontage road operations. The findings showed that wider lanes and shoulders are associated with a crash reduction on frontage roadways. Additionally, it was found that edge marking has a significant impact on safety improvement on these roadways.

Eisele et al. (2011) discussed the safety and economic impacts on conversion from two-way to one-way frontage roads. Four locations from Texas with this conversion were selected for analysis by keeping another one with no changes as the control group. CMFs were developed for these conversion works. This study developed CMFs based on fatal and injury crashes for segments and interchange intersections. Li, Lord, and Zhang (2011) evaluated the application of generalized additive models (GAMs) for CMF development in an effort to explore the safety effects of shoulder width and lane width combinations on frontage roads located in rural areas in Texas. CMFs were originated from GAMs based on data gathered on rural frontage roads in Texas. The results show that CMFs produced from GAMs were more flexible in characterizing the safety impacts of simultaneous changes in geometric and operational features than when independent CMFs were used together. Besides, the authors mentioned that increasing widths provides no safety gains. These findings were in contrast with the results in the study conducted by Lord and Bonneson (2007). Bonneson, Lord, Zimmerman, Fitzpatrick, and Pratt (2011) examined the safety impact of rural frontage-road segments in Texas by developing CMFs. In a follow-up study, Eisele, Frawley, Park, and Robertson (2012) studied eight conversion and control sites for two-way to one-way frontage road conversion. Twelve different CMFs were developed based on crash severity and crash type for frontage road conversion. Yu et al. (2021) proposed a spatially robust autoregressive model to analyze local traffic on frontage and ramp areas. The results showed the decrease in congestion propagation with increase of spatial scale. Trevino (2021) conducted a traffic noise barrier study on locations including frontage roads. The results confirmed the effectiveness of the barrier walls.

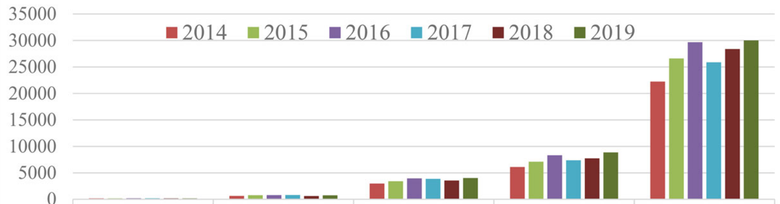
The literature review revealed that most of the frontage road-related safety studies are limited to SPF or CMF development and the determination of CMFs for two-way to one-way conversion. Specifically, in frontage road crash data analysis, no study applied cluster analysis methods such as CCA, which identifies meaningful relationships between variables in a dataset by placing them into similar groups. Crash datasets are mostly very large and difficult to interpret. Thus, to improve the interpretability of these datasets, researchers first reduce their dimensionality while ensuring that important information inherent in the data is not lost (Jolliffe & Cadima, 2016). Dimensionality reduction entails transforming the large set of variables in the dataset into a smaller one. According to the literature, conducting cluster analysis without dimensionality reduction could result in missing variables, yielding non-optimal clusters (van de Velden, D'Enza, & Palumbo, 2017). In effect, policy decisions taken based on the less accurate outcomes may not yield the expected results. A popular approach for dimensionality reduction is the Principal Component Analysis. Upon successfully reducing the dimension of the dataset, researchers then proceed to apply simple cluster analysis to identify similar and dissimilar attributes and allocate them into clusters.

The problem with applying these methods separately is that the two methods seek to optimize different criteria which may affect the quality of clusters (van de Velden et al., 2017). To solve this issue, other researchers proposed the reduced K-means (De Soete & Carroll, 1994) and factorial K-means (Yamamoto & Hwang, 2014) approaches for understanding datasets by classifying attributes into clusters. Nevertheless, due to the inability of these techniques to adequately solve the problem of missing clusters, researchers developed the CCA which combines dimension reduction and clustering methods in a joint fashion for categorical data. This novel method is advantageous in that it consistently outperforms other clustering methods that either sequentially or jointly apply dimensional reduction and clustering (van de Velden et al., 2017) and has been employed in several research in the transportation field for road safety data analysis (Das, Dutta, & Fitzpatrick, 2020; Das, Dutta, & Rahman, 2022; Das, Mahmud Hossain, et al., 2022; Rahman, Das, & Sun, 2022), airline management (Wen & Chen, 2011), and post-COVID survey data analysis (Kong, Zhang, Xiao, Das, & Zhang, 2022), among others. Another categorical data analysis method rules mining has been widely adopted in many traffic safety studies (Das, Mousavi, & Shirinzad, 2022; Das, Dutta, & Sun, 2020; Das, Dutta, & Tsapakis, 2020; Hossain, Zhou, Rahman, Das, & Sun, 2022; Tamakloe, Sam, Bencekri, Das, & Park, 2022). Due to the ability of this technique to provide meaningful and optimal clusters that are simple to visualize and understand, this study considered it in determining the key clusters of major risk factors and their combinations, which was not performed in the earlier studies.

Methodology

Data collection

The data used for this study consists of crashes that occurred on frontage roads in Texas between 2014 and 2019. As shown in Figure 1, a total of 235,522 crashes occurred on Texas roadways during 2014–2019. Although the number of crashes within the study period increased and decreased (fluctuating) in the various years within the study period, there was a general increase in the number of crashes from 32,138 in 2014 to 43,835 in 2019. It is noteworthy that the number of fatal and severe/incapacitating injury crashes peaked in 2017, with 193 and 807 crash observations, respectively. The trend reduced in 2018 but again assumed an increasing trajectory in 2019. In particular, the number of no injury cases formed the majority of the crash cases



	Fatal (K)	Incapacitating injury (A)	Non-incapacitating injury (B)	Possible injury (C)	No injury (O)
2014	167	651	2,976	6,105	22,239
2015	170	765	3,414	7,100	26,605
2016	181	794	3,957	8,313	29,669
2017	193	807	3,868	7,373	25,866
2018	167	631	3,556	7,733	28,387
2019	168	758	4,030	8,858	30,021

Figure 1. Frontage road crash counts by injury types and year.

(69.12%). Although the proportion of the total number of fatal and severe crashes from 2014 to 2019 is low (2.31%), the high economic burden associated with them is high and cannot be neglected.

Exploratory data analysis

Table 1 presents the percentage distribution of key attributes obtained from the analysis conducted in this study. It is important to note that some of the variables will implicate the information of the mainlane instead of frontage roads. A comprehensive database development for frontage roadways in Texas is out of the scope of this study. The odds ratio for crash occurrence on other road types is higher than 1 when compared with two-lane two-way main roadways. However, this association is not statistically significant. The odds ratios for crash occurrence on a curve (hillcrest), straight (grade), and straight (hillcrest) frontage roadways are higher than 1 when compared with curved (grade) section-related crashes on frontage roadways. However, only straight (hillcrest) is statistically significant. Among different traffic control devices (TCDs), the odds ratios for crash occurrences on signal light, none, and others are higher than 1 when compared with marked lanes. Signal and other TCDs are both statistically significant. For the intersection type variable, the odds ratios for intersection, intersection related (not statistically significant), non-intersection, and not reported are higher than 1 when compared to driveway access. The odds ratio for crash occurrences on U.S. and state highway frontage roadways is higher than 1 when compared to city street frontage roads. However, it is not statistically significant. The odds ratios for 5,000 vehicles per day (vpd) or less and 5,001–20,000 vpd are higher than 1 when compared to

Table 1. Percentage distribution of key attributes by crash severity.

Variable category	PDO N = 162,787	FI N = 72,735	OR	p.ratio	p.overall
Road (Main Road type)					<0.001
Two lane, two way	8220 (5.05%)	3768 (5.18%)	Ref.	Ref.	
Four or more lanes, divided	147,528 (90.6%)	65,995 (90.7%)	0.98 [0.94;1.02]	0.228	
Four or more lanes, undivided	6107 (3.75%)	2481 (3.41%)	0.89 [0.83;0.94]	<0.001	
Other road type	932 (0.57%)	491 (0.68%)	1.15 [1.02;1.29]	0.019	
Alignment					<0.001
Curve, grade	2211 (1.36%)	1063 (1.46%)	Ref.	Ref.	
Curve, hillcrest	442 (0.27%)	218 (0.30%)	1.03 [0.86;1.22]	0.776	
Curve, level	8512 (5.23%)	3567 (4.90%)	0.87 [0.80;0.95]	0.001	
Other (explain in narrative)	367 (0.23%)	168 (0.23%)	0.95 [0.78;1.16]	0.628	
Straight, grade	11,723 (7.20%)	6088 (8.37%)	1.08 [1.00;1.17]	0.057	
Straight, hillcrest	2657 (1.63%)	1458 (2.00%)	1.14 [1.04;1.26]	0.008	
Straight, level	136,589 (83.9%)	60,092 (82.6%)	0.92 [0.85;0.99]	0.019	
Unknown	286 (0.18%)	81 (0.11%)	0.59 [0.45;0.76]	<0.001	
TCD (Traffic control device)					<0.001
Marked lanes	62,716 (38.5%)	24,056 (33.1%)	Ref.	Ref.	
None	15,586 (9.57%)	6052 (8.32%)	1.01 [0.98;1.05]	0.469	
Other	21,609 (13.3%)	10,266 (14.1%)	1.24 [1.20;1.27]	0.000	
Signal light	50,152 (30.8%)	27,635 (38.0%)	1.44 [1.41;1.47]	0.000	
Yield sign	12,724 (7.82%)	4726 (6.50%)	0.97 [0.93;1.00]	0.084	
Intersection					.
Driveway access	9924 (6.10%)	3527 (4.85%)	Ref.	Ref.	
Intersection	48,625 (29.9%)	27,497 (37.8%)	1.59 [1.53;1.66]	0.000	
Intersection related	48,180 (29.6%)	19,025 (26.2%)	1.11 [1.07;1.16]	<0.001	
Non intersection	56,057 (34.4%)	22,685 (31.2%)	1.14 [1.09;1.19]	<0.001	
Not reported	1 (0.00%)	1 (0.00%)	2.81 [0.07;110]	0.524	
Road Type					<0.001
City street	1277 (0.78%)	580 (0.80%)	Ref.	Ref.	
Farm to market	4687 (2.88%)	1958 (2.69%)	0.92 [0.82;1.03]	0.142	
Interstate	90,210 (55.4%)	38,629 (53.1%)	0.94 [0.85;1.04]	0.243	
Other	845 (0.52%)	190 (0.26%)	0.50 [0.41;0.60]	<0.001	
US & state highways	65,768 (40.4%)	31,378 (43.1%)	1.05 [0.95;1.16]	0.330	
Number of Lanes					<0.001
4	48,377 (29.7%)	20,904 (28.7%)	Ref.	Ref.	
6	57,466 (35.3%)	26,034 (35.8%)	1.05 [1.03;1.07]	<0.001	
8	28,571 (17.6%)	13,209 (18.2%)	1.07 [1.04;1.10]	<0.001	
Other	28,373 (17.4%)	12,588 (17.3%)	1.03 [1.00;1.05]	0.051	
AADT (Annual Average Daily Traffic)					<0.001
> 40,000 vpd	123,923 (76.5%)	55,449 (76.7%)	Ref.	Ref.	
20,001–40,000 vpd	23,825 (14.7%)	10,118 (14.0%)	0.95 [0.93;0.97]	<0.001	
5000 vpd or less	977 (0.60%)	456 (0.63%)	1.04 [0.93;1.17]	0.458	
5001–20,000 vpd (vehicle per day)	13,163 (8.13%)	6236 (8.63%)	1.06 [1.03;1.09]	<0.001	
Population					0.000
100,000–249,999 pop.	23,040 (14.2%)	10,838 (14.9%)	Ref.	Ref.	
250,000 pop. and over	71,795 (44.2%)	37,928 (52.2%)	1.12 [1.09;1.15]	0.000	
50,000–99,999 pop	17,148 (10.6%)	6722 (9.26%)	0.83 [0.80;0.86]	0.000	
Other	30,256 (18.6%)	10,723 (14.8%)	0.75 [0.73;0.78]	0.000	
Rural	20,230 (12.5%)	6388 (8.80%)	0.67 [0.65;0.70]	0.000	
Weather					<0.001
Clear	118,024 (72.5%)	53,016 (72.9%)	Ref.	Ref.	
Cloudy	29,938 (18.4%)	13,980 (19.2%)	1.04 [1.02;1.06]	0.001	
Other	1668 (1.02%)	553 (0.76%)	0.74 [0.67;0.81]	<0.001	
Rain	13,157 (8.08%)	5186 (7.13%)	0.88 [0.85;0.91]	<0.001	
Light					<0.001
Dark, lighted	34,101 (20.9%)	16,149 (22.2%)	Ref.	Ref.	
Dark, not lighted	8122 (4.99%)	4272 (5.87%)	1.11 [1.07;1.16]	<0.001	

(*continued*)

Table 1. Continued.

Variable category	PDO N = 162,787	FI N = 72,735	OR	p.ratio	p.overall
Daylight	115,944 (71.2%)	50,464 (69.4%)	0.92 [0.90;0.94]	<0.001	
Dusk	1685 (1.04%)	743 (1.02%)	0.93 [0.85;1.02]	0.112	
Other	2935 (1.80%)	1107 (1.52%)	0.80 [0.74;0.86]	<0.001	
Collision					0.000
Angle - both going straight	21,066 (12.9%)	17,128 (23.5%)	Ref.	Ref.	
Other	65,963 (40.5%)	27,119 (37.3%)	0.51 [0.49;0.52]	0.000	
SD both straight-rear end	20,937 (12.9%)	9467 (13.0%)	0.56 [0.54;0.57]	0.000	
SD both straight-sideswipe	22,330 (13.7%)	4037 (5.55%)	0.22 [0.21;0.23]	0.000	
SD one straight-one stopped	32,491 (20.0%)	14,984 (20.6%)	0.57 [0.55;0.58]	0.000	
Other					0.000
Attention diverted	15,433 (9.48%)	8277 (11.4%)	Ref.	Ref.	
Not applicable	63,195 (38.8%)	33,041 (45.4%)	0.97 [0.95;1.00]	0.095	
Other	48,864 (30.0%)	19,952 (27.4%)	0.76 [0.74;0.79]	0.000	
Slowing/stopping	18,081 (11.1%)	7950 (10.9%)	0.82 [0.79;0.85]	0.000	
Vehicle changing lanes	17,214 (10.6%)	3515 (4.83%)	0.38 [0.36;0.40]	0.000	

Note: PDO: no injury/property damage only; FI: Fatal and injury; OR: odds ratios.

roadways with more than 40,000 vpd traffic volume. However, only 5,001–20,000 vpd roadway is statistically significant. This traffic volume measure is needed to be interpreted carefully as the current study used overall AADT measures on the main roadways. The odds ratio for roadways under the area with 250,000 or more population is higher than 1 when compared to frontage roadways under the area with a population between 100,000 to 249,999. This association is statistically significant. The odds ratio for cloudy day crashes on frontage roads is higher than 1 when compared to clear day crashes on frontage roads, which is statistically significant. The odds ratio of crashes on frontage roadways with dark (not lighted) is higher than 1 when compared to crashes on frontage roadways with dark (lighted).

Cluster correspondence analysis

As described earlier, CCA was applied to this study due to a large number of categorical variables associated with the frontage roadways crashes. It can be observed in Table 1 that there are 12 variables, which make more than fifty categories or attributes. For such a large number of categorical variables, CCA is an appropriate method to identify the similarity patterns. CCA applies correspondence analysis to the cross-tabulation of the cluster membership and the variable attributes. This process confirms the variance among clusters as maximum and the categories with different distributions among clusters to be optimally separated. CCA was applied on the contingency matrix (the cluster by categories contingency matrix) to allocate clusters and K-means was applied to reduce the space coordinates (obtained from the CCA interactions). A very short overview on the theoretical framework is described below, which is based on van de Velden et al. (2017). Readers can consult van de Velden et al. (2017) for comprehensive overview.

Consider the frontage road-related crash dataset consists of n observations such that each observation is associated with a set of p categorical variables. Then, the frontage crash dataset can be represented as a matrix $\boldsymbol{Z}_{n,Q}$ where $Q = \sum_{j=1}^{p} q_j$. In this matrix representation, the rows and columns denote the crash observations and categories, respectively. If K connotes the number of clusters, then, membership to a cluster can be coded and represented as the $n \times K$ indicator matrix $\boldsymbol{Z}_K$ with dimension d. By employing the CCA approach, optimal scaling row values representing clusters, and columns representing the individual categories with a maximum between cluster variance are obtained. Based on the distributions over the categories, the clusters are optimally separated. Given that $M = I_n - 1_n 1'_n/n$ and D_z represents the diagonal matrix with the assumption that $D_z 1_Q = Z' 1_n$, the objective function for the CCA technique can be defined as follows (Rahman et al., 2022; van de Velden et al., 2017):

$$max\emptyset_{clusca}(Z_K, B^*) = \frac{1}{p} traceB^{*'} D_z{}^{-1/2} Z' M Z_K D_z{}^{-1} Z'_K M Z D_z{}^{-1/2} B^* \tag{1}$$

When B^* is fixed, the optimization problem becomes a K-means clusting issue. The resulting K-means objective function defined below is then solved by maximizing $\emptyset(Z_K, B^*)$ with respect to Z_K (van de Velden et al., 2017):

$$min\emptyset'_{clusca}(Z_K, G) = |\sqrt{\frac{n}{p}} MZD_z{}^{-\frac{1}{2}} B^* - Z_K G|^2 \tag{2}$$

where G represents the matrix with cluster average measures, B denotes a column coordinate matrix with a rank of k. The key steps are listed below:

- Randomly assign objects to clusters to get the initial cluster allocation, obtain the initial solution by utilizing MCA
- Apply CCA to the contingency matrix to obtain the category quantifications B (column coordinates) and cluster mean
- Get the coordinates for the objects using category quantification
- Apply K-means
- Iterate the second step to the fourth step until convergence remains constant.

The CCA method used in this paper was run in R software (Markos, D'Enza, & van de Velden, 2019). Prior to applying this method, the researcher must specify an appropriate number of clusters and dimensionality. This process needs careful consideration to aid better cluster separation and facilitate the extraction of meaningful clusters. To evaluate these

Table 2. Centroids and size of the clusters.

Cluster	Fatal and injury crashes			No injury crashes		
	Dim 1	Dim 2	Size	Dim 1	Dim 2	Size
Cluster 1	−0.00278	−0.00196	26,652	−0.00036	−0.00040	52,538
Cluster 2	−0.00004	0.00343	21,748	−0.00001	0.00239	48,396
Cluster 3	0.00152	0.00060	13,660	−0.00226	−0.00167	31,942
Cluster 4	0.00507	−0.00286	10,675	0.00306	−0.00138	29,911

hyperparameters, the function '*tuneclass*' in the R software package known as '*clustrd*' was employed. The function applies well-established distance-based statistics, the average silhouette width index, and the Calinski-Harabasz index, to assess the quality of the clustering solutions in deciding on the appropriate number of clusters and dimensionality. The output of the parameter tuning function provides the optimal number of clusters and dimensionality. For a detailed description of the CCA method including the parameter tuning process used in this study, interested readers are directed to the joint dimension reduction and clustering papers (Markos et al., 2019; Rahman et al., 2022; van de Velden et al., 2017).

Results and discussions

This study applied CCA on two frontage road crash datasets based on injury types: fatal and injury (FI) and no injury (PDO). In the initial stage, the categorical values are randomly assigned to the optimized number of clusters using MCA. The ultimate target is to make the convergence constant to determine the locations of the variables in each cluster. Table 2 lists the centroid location of the selected clusters in each dataset. The count of cluster selection was based on the optimization of the criterion measures. The criterion measures for these two datasets are 4.367 (fatal and injury) and 4.172 (no injury). Upon conducting several trials, it was established that a two-dimensional, four-cluster solution was more appropriate. Thus, for each dataset, four clusters were finally selected. Table 2 shows that the first two clusters of each dataset contain more than 62% of the data. The following section provides details on the results from each dataset.

Fatal and injury crashes

From the plot shown in Figure 2, Cluster 1 (36.6%), which relates primarily to the driver being at an intersection, attention being diverted from driving, collision angles with both going straight, and traffic signal lights and other TCDs being used, resides mostly in the second and third quadrants. Cluster 2 (29.9%), which related primarily to int. non-intersection (segment crashes) were either marked lanes or no traffic control devices are used, with straight-rear end or straight side-sweep collisions, one of the drivers' changing lanes,

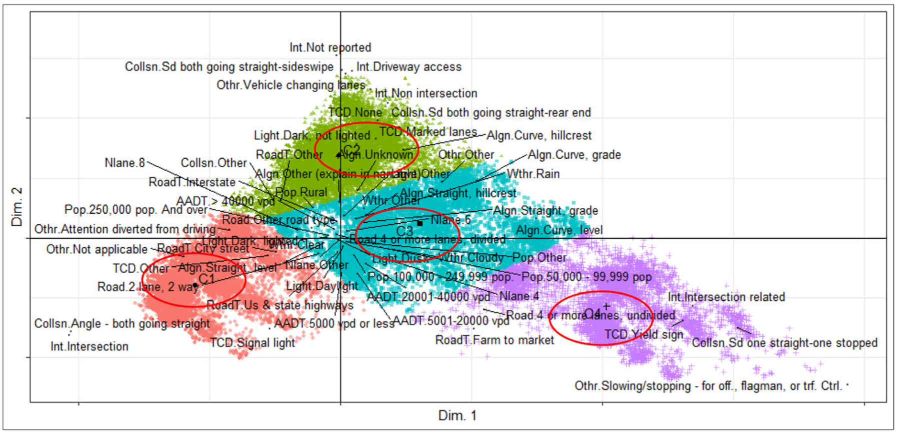


Figure 2. Clusters developed from fatal and injury crash data.

when there is driveway access, and when it is dark with no lighting, resides within the first and second quadrants. Cluster 3 (18.8%), which primarily relates to int. intersection collisions, collisions with one stopped and one straight or with both going straight (a rear end), either a yield sign being used as a traffic control device or none are used, a level curve or graded curve, and lighting conditions being dark, resides mostly in the first and fourth quadrant. Cluster 4 (14.7%), which primarily relates to slowing/stopping, crashes occurring at an intersection, collisions with one straight and one stopped, traffic control devices of yield signs or signal lights being used, at undivided roads of four or more lanes, and on farm to market roads, and with curve align, resides mostly in the fourth quadrant. Cluster 3 is the closest to the center.

Cluster 1 (C1; Figure 3)—Distraction related crashes at signalized intersections

Cluster 1 (C1; Figure 3) indicates crashes at signalized intersections. This cluster represents around 37% of the data. These crashes often occur when attention is diverted from driving, and the collision occurs when both vehicles are going straight. The literature identified that not obeying traffic rules, which includes inattentive driving, leads to fatal crashes on frontal roads (Mallipaddi & Anderson, 2020). Traffic signal lights are also associated with this type of collision, along with other types of traffic control devices. This finding makes sense, as drivers tend to rush to join freeways without adequately checking for oncoming vehicles, particularly when the yellow light turns on. This act is likely to cause serious interactions with oncoming vehicles on the freeway, leading to fatalities.

Cluster 2 (C2; Figure 3)—Segment related crashes at dark unlighted conditions

In Cluster 2 (C2), the non-intersection (i.e., segment) has the longest bar on the positive side, indicating that it has the strongest association to other

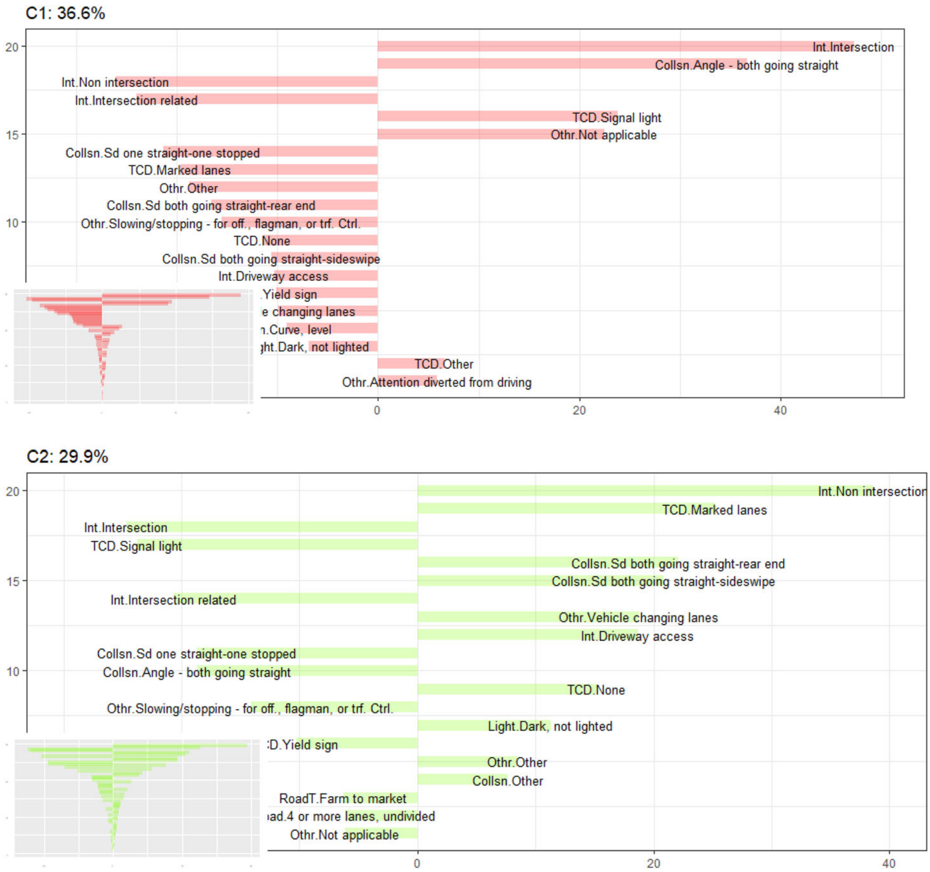


Figure 3. Cluster 1 and cluster 2 (fatal and injury crash data).

categories in this cluster. The graph shows that collisions of this sort often happen when there are marked lanes or no traffic control devices (TCDs), and the most common collisions to arise in this scenario are straight-rear end or straight side-sweep collisions. Other conditions associated with this cluster are when one of the drivers changes lanes when there is driveway access and when it is dark with no lighting. This finding is in line with the literature; it was identified that crashes on frontage roads are likely to be severe during the night. It is noteworthy that the authors did not specify the lighting condition present at the time of the crash during the night (Mallipaddi & Anderson, 2020).

Cluster 3 (C3; Figure 4)—Yield signed intersection locations and segments with no TCDs

In Cluster 3 (C3), intersection related had the longest bar on the positive side of C3, indicating it had the strongest association with other categories in this cluster, such as collisions with one stopped and one straight or with

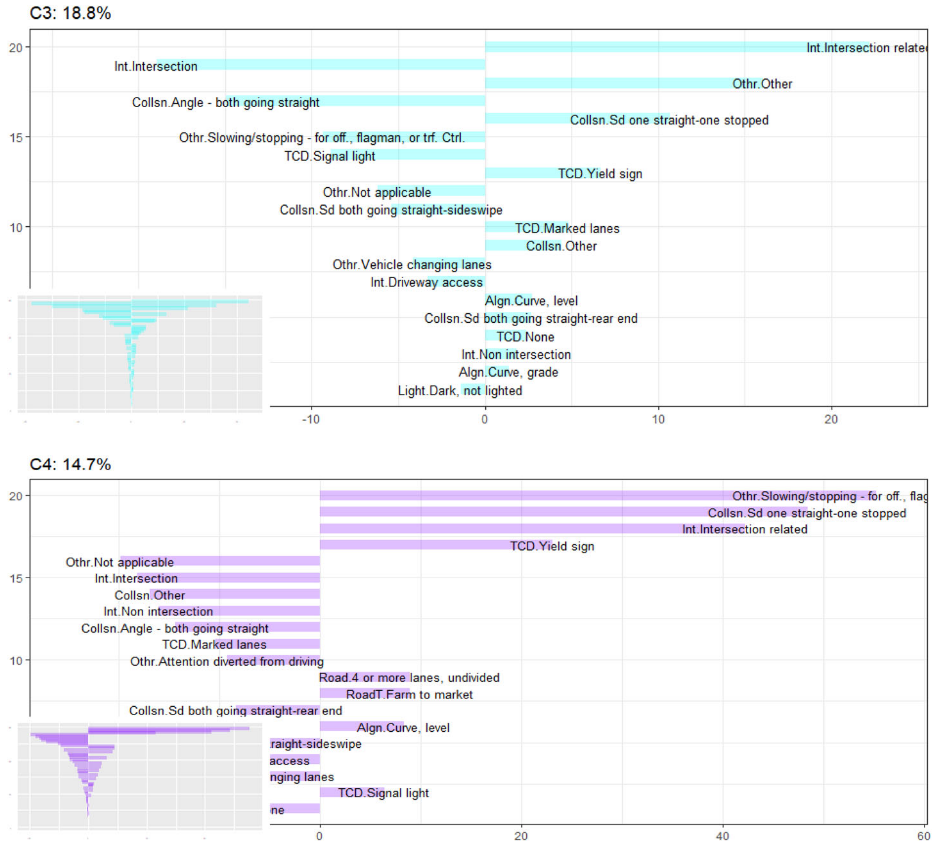


Figure 4. Cluster 3 and cluster 4 (fatal and injury crash data).

both going straight (a rear-end). Further, the other group is associated with segments with no TCDs. These crashes are associated with segments with either a level curve or a graded curve. In line with previous studies, segments with yield signs are associated with fatal crashes (Zhao et al., 2018). Besides, in the literature, intersections have been noted as areas where fatal crashes are likely due to the complexity and the high number of interactions among vehicles at these sections (Tay & Rifaat, 2007). Thus, drivers on frontal roads that are not cautious could be involved in fatal or injury crashes at intersection locations.

Cluster 4 (C4, Figure 4)—Intersection crashes on undivided roadways

In cluster 4 (C4), slowing/stopping had the longest bar on the positive side of C4. This indicates it has the strongest association with other categories in this cluster, such as occurring at an intersection, collisions with one straight and one stopped, with traffic control devices of yield signs or signal lights being used, on undivided roads of four or more lanes, on farm to market roads, and with aligned curvatures. Intuitively, inattentive, or

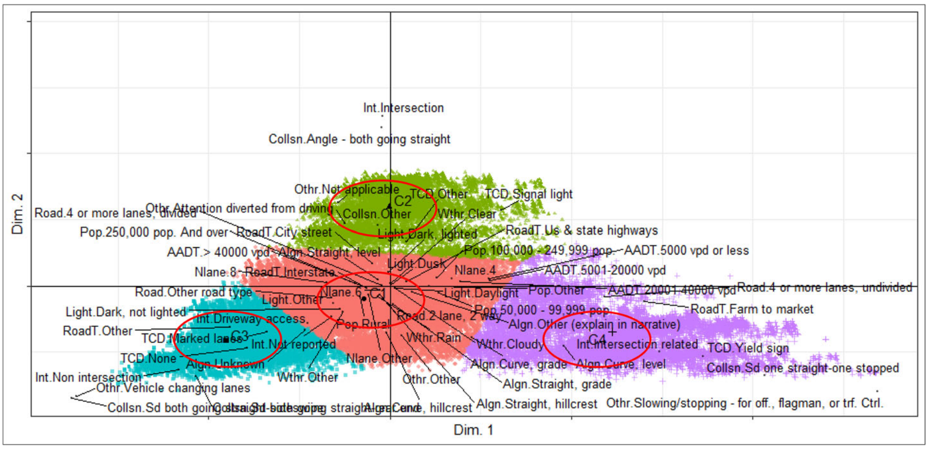


Figure 5. Clusters developed from no injury crash data.

confused drivers are likely to be involved in fatal or injury crashes when they fail to yield or obey signal lights, as identified in the literature (Zhao et al., 2018). This finding highlights the need to educate drivers on the need to obey roadway signs and signals, particularly on frontage roads.

No injury crashes

From the plot shown in Figure 5, Cluster 1 (32.3%), which is primarily related to driveway access, marked lanes being used or no TCDs being used, segment areas, crashes occurring in the rain, crashes occurring in the dark with no lighting, and crashes at a curve, grade, is almost equally in all four quadrants (although a little more is in quadrant three and four). Cluster 2 (29.7%), which is primarily related to int. intersection, collision angle with both going straight, traffic control devices of either a signal light or other unlisted types being used, and driver attention being diverted from driving, is in quadrant one and two. Cluster 3 (19.6%), which is primarily related to collisions with both drivers going straight, the other vehicle changing lanes, crashes occurring at segments, collisions with both going resulting in straight rear-ends, there being either marked lanes or no TCDs, and 8 main lanes, is located mostly in quadrant three. Cluster 4 (18.4%), which is primarily related to collisions with one driver straight and one stopped, slowing/stopping, crashes occurring at an intersection, yield signs or traffic lights being present, crashes on curve, level, crashes on farm to market roads, and crashes on undivided roads with four or more lanes, is mostly located in quadrant four. Cluster one is located closest to the center of the graph.

Cluster 1 (C1; Figure 6)—Segment crashes in dark conditions and rain

In cluster 1, driveway access has the longest bar on the positive side, indicating that it had the strongest association with other categories in this

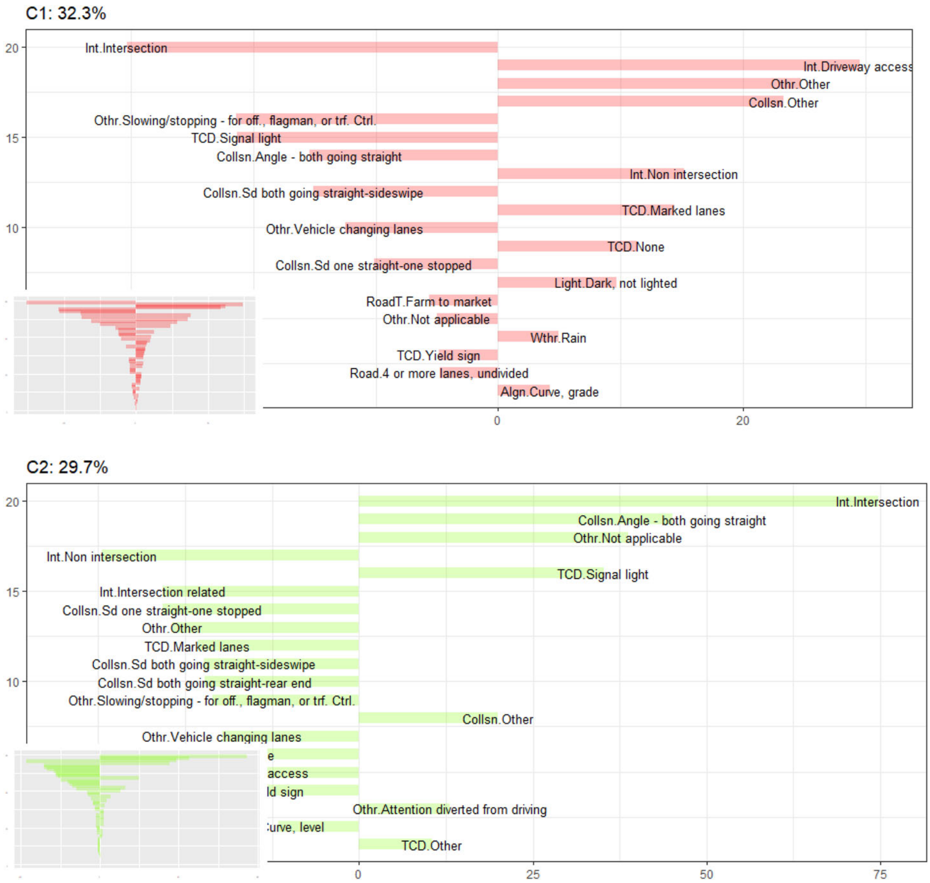


Figure 6. Cluster 1 and cluster 2 (no injury crash data).

cluster, such as either marked lanes being used or no TCDs being used, collisions occurring at segment areas, collisions occurring in the rain, collisions occurring in the dark with no lighting, and at a curve, grade. These findings are plausible as drivers are likely to be more careful as they drive on dark-unlighted roads with curves/grades. Similar results were identified in previous safety research, which showed that fatal crashes are more likely when drivers cruise on straight road segments due to the increased chance of relaxing on such roads as their perception of being safe increases (Tamakloe, Hong, & Park, 2020).

Cluster 2 (C2; Figure 6)—Crashes at signalized intersection with both drivers going straight

In cluster 2, int. intersection has the longest bar on the positive side, indicating it has the strongest association with other categories in this cluster, such as a collision angle with both going straight, traffic control devices of

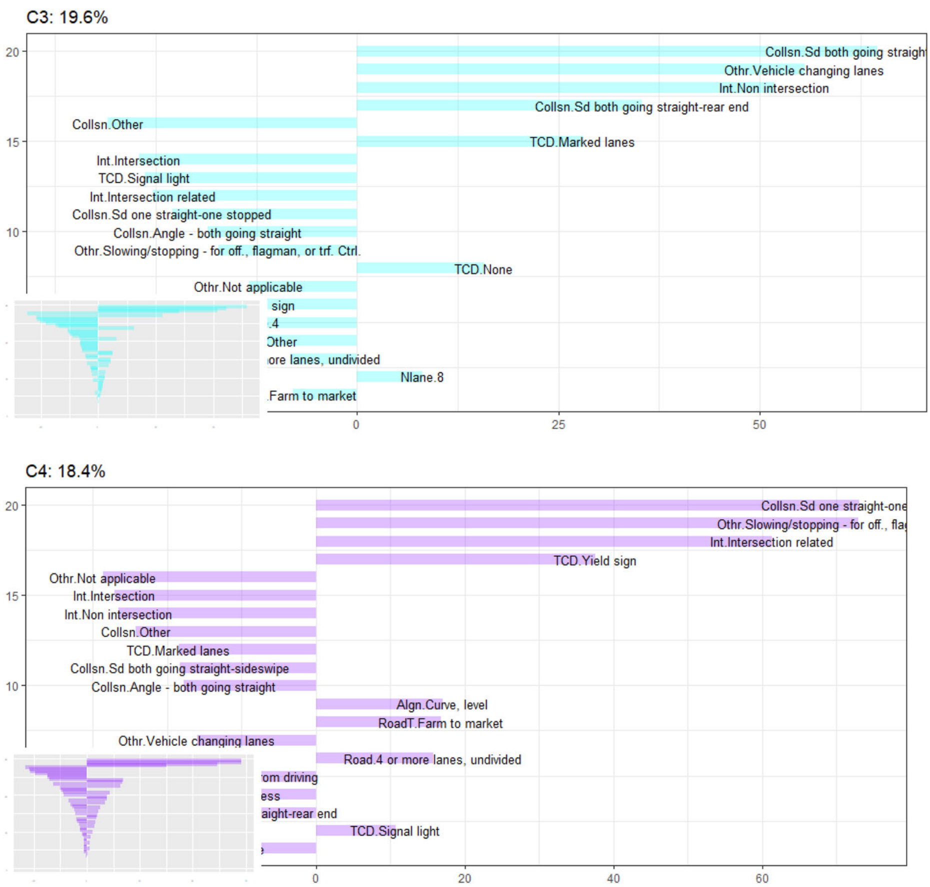


Figure 7. Cluster 3 and cluster 4 (non-injury crash data).

either a signal light or other unlisted types being used, and driver attention being diverted from driving. Signalized intersections are highly controlled areas. Intuitively, drivers making straight maneuvers in such areas are less likely to have fatal injuries as the chance of sideswipes are higher (Ma, Zhao, Chien, & Dong, 2015).

Cluster 3 (C3; Figure 7)—Segment crashes with both drivers going straight with marked lanes or no TCDs

In cluster 3, collisions with both drivers going straight have the longest bar on the positive side, indicating it has the strongest association with other categories in this cluster, such as the other vehicle changing lanes, collisions occurring at segments, collisions with both going into straight rear-ends, roads with either marked lanes or no TCDs, and roads with 8 main lanes on the major facilities. Improper lane change maneuvers are likely to result in sideswipes in vehicles cruising in the same direction. As indicated in the literature, sideswipe crashes are less likely to result in fatal injury (Ma et al., 2015).

Cluster 4 (C4; Figure 7)—Intersection related collisions on undivided roadways

In cluster 4, collisions with one driver straight and one stopped has the longest bar on the positive side, indicating it has the strongest association with other categories in this cluster, such as slowing/stopping, collisions occurring at an intersection, yield signs or traffic lights being present, on curve, level, on farm to market roads, and on undivided roads with four or more lanes. Drivers are more likely to slow down on approaching curves. Besides, they are likely to be even more cautious when no traffic signals are present on undivided road segments. Due to the reduced speeds and increased attentiveness at such segments, it is less likely to have fatal crashes. In the literature, median-divided segments have been noted to be crash-prone. These segments are likely to have large-sized and severe crashes and are plausible as medians are installed on segments with high-speed limits (Pande & Abdel-Aty, 2009; Tamakloe, Lim, Sam, Park, & Park, 2021). There is a high chance that drivers at such segments would be caught unaware of the traffic ahead, leading to crashes with fatal consequences.

In total, four clusters were identified for each severity level. Regarding fatal and injury crashes, it was identified that signalized inattentive driving intersections (C1), dark-unlighted non-intersection segments (C2), yield-signed intersections, and those with no TCD's (C3), and undivided road intersections (C4) were identified as critical risk factors. Clearly, intersections were identified as critical areas prone to fatal or injury crashes as in previous studies. Besides, the kind of control at the intersection also showed up as a critical risk factor leading to fatal or injury crashes at frontal roads in Texas. Moreover, fatal/injury crashes were likely to occur in dark conditions, consistent with the literature. Concerning no injury crashes, it was identified that dark conditions and rain (C1), straight maneuvers at signalized intersections (C2), straight maneuvers on marked lanes (C3), and undivided road intersection segments (C4) were identified as critical risk factors. In this data category, signalized intersections were identified as key locations for no injury crashes. Additionally, the key factors included straight maneuvers.

Conclusions

Although frontage roads have been a design solution in terms of providing access to freeways from adjacent locations, they have become common areas for fatal crashes due to the changes in the land use near them. Nevertheless, researchers have paid less attention to analyzing the risk factors influencing these crashes, and the few studies that explored frontage road crashes did not explore the patterns associated with key variable categories or attributes influencing such crashes. As this approach is robust in terms of unearthing clusters of variables or attributes influencing crashes and provides visual guidance for easily understanding the results, we posited that it would provide valuable outcomes for use by safety practitioners worldwide. Thus, a CCA approach was used in this study to allocate crash-risk factors contributing to frontage road crashes into meaningful groups. To understand how these factors differ by severity level, the study segregated the crash data set into two groups based on the severity level of the crash – fatal and injury, and no injury. Four clusters in each severity level provided several risk clusters. These risk clusters can further be analyzed to make data-driven decision making.

This study is expected to provide valuable information to guide the formulation of relevant policies to help reduce the severity of crashes on frontage roads. Although this study used crash data from Texas, the results can be extended to other frontage roads with similar traffic conditions. As with every study, this analysis has a few limitations. Although the methodology used in this study provides an approach for clustering risk factors into meaningful groups, it does not provide associations between the factors in each group. In the future, there is the need to expand this analysis by using more robust data mining tools such as Bayesian latent class model (Das, 2022; Das, Mahmud Hossain, et al., 2020), empirical Bayes geometric mean (Das, Bibeka, Sun, "Tracy" Zhou, & Jalayer, 2019), quantitative risk analysis (Weng, Gan, & Zhang, 2021), simulation studies (Gan, Weng, & Zhang, 2021) which can discover the relationships existing between various risk factors influencing crashes on frontal roads.

Acknowledgment

The authors like to thank the anonymous reviewers for their thoughtful comments and suggestions.

Declaration of conflicting interest

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

ORCID

Subasish Das http://orcid.org/0000-0002-1671-2753
Reuben Tamakloe http://orcid.org/0000-0002-9605-8439
Boniphace Kutela http://orcid.org/0000-0002-5450-1623
Ahmed Hossain http://orcid.org/0000-0003-1566-3993

References

Bhat, C. R., Born, K., Sidharthan, R., & Bhat, P. C. (2014). A count data model with endogenous covariates: Formulation and application to roadway crash frequency at intersections. *Analytic Methods in Accident Research*, *1*, 53–71. doi:10.1016/j.amar.2013.10.001

Bonneson, J., Lord, D., Zimmerman, K., Fitzpatrick, K., & Pratt, M. (2011). Calibration of accident modification factors for rural frontage road segments. *Accident Reconstruction Journal*, *21*(2), 53–57.

Daoud, J. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series*, *949*, 012009. doi:10.1088/1742-6596/949/1/012009

Das, S. (2022). *Aritificial intelligence in highway safety* (1st ed.). Boca Raton, FL: CRC Press.

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. doi:10.1016/j.aap.2017.11.016

Das, S., Bibeka, A., Sun, X., "Tracy" Zhou, H., & Jalayer, M. (2019). Elderly pedestrian fatal crash-related contributing factors: Applying empirical Bayes geometric mean method. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(8), 254–263. doi:10.1177/0361198119841570

Das, S., Chatterjee, S., Mitra, S., Singh, V., Asari, V. K., & Kumar, S. (2021). Improper passing and lane-change related crashes: Pattern recognition using association rules negative binomial mining. In R. B. Patel (Ed.), *Computational methods and data engineering, advances in intelligent systems and computing* (pp. 561–575). Berlin, Germany: Springer.

Das, S., Dutta, A., & Fitzpatrick, K. (2020b). Technological perception on autonomous vehicles: Perspectives of the non-motorists. *Technology Analysis & Strategic Management*, *32*(11), 1335–1352. doi:10.1080/09537325.2020.1768235

Das, S., Mousavi, S., & Shirinzad, M. (2022). Pattern recognition in speeding related motorcycle crashes. *Journal of Transportation Safety & Security*, *14*(7), 1121–1138. doi:10.1080/19439962.2021.1877228

Das, S., Dutta, A., & Rahman, M. A. (2022b). Pattern recognition from light delivery vehicle crash characteristics. *Journal of Transportation Safety & Security*, 1–19. doi:10.1080/19439962.2021.1995800

Das, S., Dutta, A., & Tsapakis, I. (2020a). Automated vehicle collisions in California: Applying Bayesian latent class model. *IATSS Research*, *44*(4), 300–308. doi:10.1016/j.iatssr.2020.03.001

Das, S., Dutta, A., & Sun, X. (2020c). Patterns of rainy weather crashes: Applying rules mining. *Journal of Transportation Safety & Security*, *12*(9), 1083–1105. doi:10.1080/19439962.2019.1572681

Das, S., Mahmud Hossain, M., Ashifur Rahman, M., Kong, X., Sun, X., & Al Mamun, G. M. (2022a). Understanding patterns of moped and seated motor scooter (50 Cc or less) involved fatal crashes using cluster correspondence analysis. *Transportmetrica A: Transport Science*, 1–21.

De Soete, G., & Carroll, J. D. (1994). K-means clustering in a low-dimensional Euclidean space. In *Studies in Classification, Data Analysis, and Knowledge Organization* (pp. 212–219). Berlin, Germany: Springer. doi:10.1007/978-3-642-51175-2_24

Eisele, W. L., Frawley, W. E., Park, E. S., & Robertson, J. (2012). Safety And Economic Impacts Of Converting Two-Way Frontage Roads To One-Way Operation.

Transportation Research Record: Journal of the Transportation Research Board, *2301*(1), 66–75. doi:10.3141/2301-08

Eisele, W. L., Yager, C. E., Brewer, M. A., Frawley, W. E., Park, E. S., Lord, D., Robertson, J. A., & Kuo, P. (2011). *Safety and economic impacts of converting two-way frontage roads to one-way: Methodology and findings*. Austin, TX: Texas Department of Transportation.

Gan, X., Weng, J., & Zhang, J. (2021). Evaluation of travel delay and accident risk at moving work zones. *Journal of Transportation Safety & Security*, *13*(6), 622–641. doi:10.1080/19439962.2019.1659464

Gattis, J. L., Hanning, C., & Duncan, L. K. (2008). Effects of Frontage Road Conversion. Report MBTC 2073. Mack–Blackwell Rural Transportation Center, University of Arkansas, Fayetteville.

Hossain, M. M., Zhou, H., Rahman, M. A., Das, S., & Sun, X. (2022). Cellphone-distracted crashes of novice teen drivers: Understanding associations of contributing factors for crash severity levels and cellphone usage types. *Traffic Injury Prevention*, *23*(7), 390–397. doi:10.1080/15389588.2022.2097667

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *374*(2065), 20150202. 10.1098/rsta.2015.0202.

Kockelman, K., Machemehl, R., Overman, A. W., Madi, M., Sesker, J., Peterman, J., & Handy, S. (2021). *Frontage roads in Texas: A comprehensive assessment*. Center for Transportation Research at The University of Texas at Austin. http://www.utexas.edu/research/ctr/pdf_reports/1873_2.pdf

Kong, X., Zhang, A., Xiao, X., Das, S., & Zhang, Y. (2022). Work from home in the post-COVID world. *Case Studies on Transport Policy*, *10*(2), 1118–1131. doi:10.1016/j.cstp.2022.04.002

Kusselson, S. B. (2013). *Investigating how land use patterns affect traffic accident rates near frontage road cross-sections: A case study on interstate 610 in Houston, Texas*. Stillwater, OK: Oklahoma State University.

Kutela, B. (2022). The role of crosswalk-related features on drivers' spatial yielding compliance at signalized midblock crosswalks. *Journal of Traffic and Transportation Engineering (English Edition)*, *9*(4), 666–675. doi:10.1016/j.jtte.2021.11.001

Li, X., Lord, D., & Zhang, Y. (2011). Development of accident modification factors for rural frontage road segments in texas using generalized additive models. *Journal of Transportation Engineering*, *137*(1), 74–83. doi:10.1061/(ASCE)TE.1943-5436.0000202

Lord, D., & Bonneson, J. A. (2007). Development of accident modification factors for rural frontage road segments in Texas. *Transportation Research Record: Journal of the Transportation Research Board*, *2023*(1), 20–27. doi:10.3141/2023-03

Ma, Z., Zhao, W., Chien, S. I.-J., & Dong, C. (2015). Exploring factors contributing to crash injury severity on rural two-lane highways. *Journal of Safety Research*, *55*, 171–176. doi:10.1016/j.jsr.2015.09.003

Machsus, Prayogo, I., Hayati, D. W., & Utanaka, A. (2017). Road safety analysis on Achmad Yani Frontage Road Surabaya. *IOP Conference Series: Materials Science and Engineering*, *267*, 012027. doi:10.1088/1757-899X/267/1/012027

Mallipaddi, V., & Anderson, M. (2020). Analysis of crashes on freeway weaving sections. Presented at International Conference on Transportation and Development 2020, Seattle, Washington, May 26–29, 2020.

Markos, A., D'Enza, A. I., & van de Velden, M. (2019). Beyond tandem analysis: Joint dimension reduction and clustering in R. *Journal of Statistical Software*, *91*(10), 1–24. doi:10.18637/jss.v091.i10

Pande, A., & Abdel-Aty, M. (2009). A novel approach for analyzing severe crash patterns on multilane highways. *Accident Analysis & Prevention*, *41*(5), 985–994. doi:10.1016/j.aap.2009.06.003

Rahman, M. A., Das, S., & Sun, X. (2022). Using cluster correspondence analysis to explore rainy weather crashes in Louisiana. *Journal of the Transportation Research Board*. *2676*(8), 159–173.

Tamakloe, R., Das, S., Nimako Aidoo, E., & Park, D. (2022). Factors affecting motorcycle crash casualty severity at signalized and non-signalized intersections in Ghana: Insights from a data mining and binary logit regression approach. *Accident Analysis & Prevention*, *165*(February), 106517. doi:10.1016/j.aap.2021.106517

Tamakloe, R., Hong, J., & Park, D. (2020). A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data. *Accident Analysis & Prevention*, *146*, 105736. doi:10.1016/j.aap.2020.105736

Tamakloe, R., Lim, S., Sam, E. F., Park, S. H., & Park, D. (2021). Investigating factors affecting bus/minibus accident severity in a developing country for different subgroup datasets characterised by time, pavement, and light conditions. *Accident Analysis & Prevention*, *159*, 106268. doi:10.1016/j.aap.2021.106268

Tamakloe, R., Sam, E. F., Bencekri, M., Das, S., & Park, D. (2022). Mining groups of factors influencing bus/minibus crash severities on poor pavement condition roads considering different lighting status. *Traffic Injury Prevention*, *23*(5), 308–314. doi:10.1080/15389588.2022.2066658

Tay, R., & Rifaat, S. M. (2007). Factors contributing to the severity of intersection crashes. *Journal of Advanced Transportation*, *41*(3), 245–265. doi:10.1002/atr.5670410303

Trevino, M. (2021). *Traffic noise barrier study on SH 190 in Texas*. Austin, TX: Texas Department of Transportation.

van de Velden, M., D'Enza, A. I., & Palumbo, F. (2017). Cluster correspondence analysis. *Psychometrika*, *82*(1), 158–185. doi:10.1007/s11336-016-9514-0

Wen, C.-H., & Chen, W.-Y. (2011). Using multiple correspondence cluster analysis to map the competitive position of airlines. *Journal of Air Transport Management*, *17*(5), 302–304. doi:10.1016/j.jairtraman.2011.03.006

Weng, J., Gan, X., & Zhang, Z. (2021). A quantitative risk assessment model for evaluating hazmat transportation accident risk. *Safety Science*, *137*, 105198. doi:10.1016/j.ssci.2021.105198

Xu, J., Kockelman, K., & Wang, Y. (2014). Modeling crash and fatality counts along mainlanes and frontage roads across Texas: The roles of design, the built environment, and weather. Presented at the 93rd Transportation Research Board Annual Meeting, Washington, D.C.

Yamamoto, M., & Hwang, H. (2014). A general formulation of cluster analysis with dimension reduction and subspace separation. *Behaviormetrika*, *41*(1), 115–129. doi:10.2333/bhmk.41.115

Yu, H., Deng, S., Lu, C., Tang, Y., Yu, S., Liu, L., & Ji, T. (2021). Research on the evolution mechanism of congestion in the entrances and exits of parking facilities based on the improved spatial autoregressive model. *Journal of Advanced Transportation*, *2021*, e8380247. doi:10.1155/2021/8380247

Zhang, Y., Xie, Y., & Li, L. (2012). Crash frequency analysis of different types of urban roadway segments using generalized additive model. *Journal of Safety Research*, *43*(2), 107–114. doi:10.1016/j.jsr.2012.01.003

Zhao, Q., Goodman, T., Azimi, M., & Qi, Y. (2018). Roadway-related truck crash risk analysis: Case studies in Texas. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(34), 20–28. doi:10.1177/0361198118794055