

Research

Multivariate analysis of pedestrian-related crashes on Vienna's roads

Radmila Magusic¹

Received: 1 November 2023 / Accepted: 9 January 2024

Published online: 23 January 2024

© The Author(s) 2024 [OPEN](#)

Abstract

Pedestrians are the least protected traffic group when compared to other participants. When compared to moto riders (independent of engine power) they have head protection provided by helmet or vehicle metal shield along with airbag that is providing highest level of body protection for personal vehicles drivers. This vulnerable group is characterized by the high presence of very young participants whose psychological characteristics place them in a vulnerable risky subgroup. According to the police original data on crashes involving pedestrians in the town of Vienna, since 2010 there was highest number of crashes in 2012 and after slow decrease is recorded but still high in total number of crashes involving only pedestrians is an extremely worrying problem.

This research is essential to address leading characterizations in crashes with the aim to answer what is current trend in crash occurrence during 2010–2020 inside Vienna municipality, and what is predicted trend. Is there significant and distinctive difference based on gender and age with specific conditions under which crashes are occurring influencing different injury degree. Multiple regression undoubtedly points fields for action in statistically based findings providing the most important answer to this research: why crashes do occur so frequently and what is leading cause of injured pedestrians. Stepwise procedure in discriminant analysis at statistically significant level shows what differentiates injured and not injured pedestrians.

Highlights

- Highest increase in percentage of variance explained is causing children involvement, person behaviour, police assumption on main cause of crash and lane
- Road conditions, crash type, gender, age, child involvement and “hit and run” crash are statistically significantly predicting pedestrians' injuries
- Crashes that predominately result in more severe injuries involve foreign older pedestrians and child with mark “on the way to school” when driver left crash spot

Keywords Road safety · Pedestrians · Crash · Factors · Predictors · Regression

✉ Radmila Magusic, radica.magusic@gmail.com | ¹Svetosimunska Street 84, 10 000 Zagreb, Croatia



Discover Analytics

(2024) 2:3

| <https://doi.org/10.1007/s44257-024-00009-y>

1 Introduction

Road crashes are not rarely expressed through injuries and costs they cause. The most recent road safety report for Austria shows the cost of road crashes in year 2016 was 3.3% of GDP or 9.7 billion euros. Leading cause of fatal crashes was speed, followed by alcohol and drugs intoxicated participants [1].

During eleven years of observed period, from 2010 until 2020, in the city centre of Vienna, frequency of crash frequencies has highest number in year 2012 (N = 1520), as visible from Fig. 1. In the following years total number slowly decreases but 1000 crashes involving pedestrians per year is worrying number and still far from 50% reduction defined in National strategy [2]. When total number of crashes involving pedestrians is observed through decrease in chain index, then in all ten years negative decrease was recorded until Year 2012 and in 2016 while chain index in all other observed years showed increase. Impairment in dataset is spanning from fatigue, medication, health and arousal, but the leading ones are alcohol and drugs. Further observing of defined impairment in light, severe and death injuries again reveal alcohol and drugs as leading impairments. Along eleven years in observed period, total number of crashes is slowly decreasing in Vienna and at national level as well which indicates continuum in national dedication towards safer road environment for all its participants. National road safety program implemented in year 2011 shows efficiency of strategy and measures but question that stays unanswered is whether decrease will continue in future, or it will stay at some point. Due to data, it seems it has stopped at 1000 pedestrian related crashes per year. However, not sufficient attention is dedicated in understanding what is causing crashes every year and it presents motivation to engage in this research.

2 Methodological approach

Main goal of this research is to determine relevant differences based on gender and age, and influential predictors on decrease in pedestrian-related crashes in Vienna. The main problem of research is defined by the question: is there statistically significant difference based on gender, children involvement and young vs older pedestrians when compared to other age groups? To answer the research problem posed, the following research questions are defined:

1. Are there notable and distinctive differences in terms of gender and age, as well as specific circumstances surrounding crashes involving pedestrians that impact the severity of injuries?

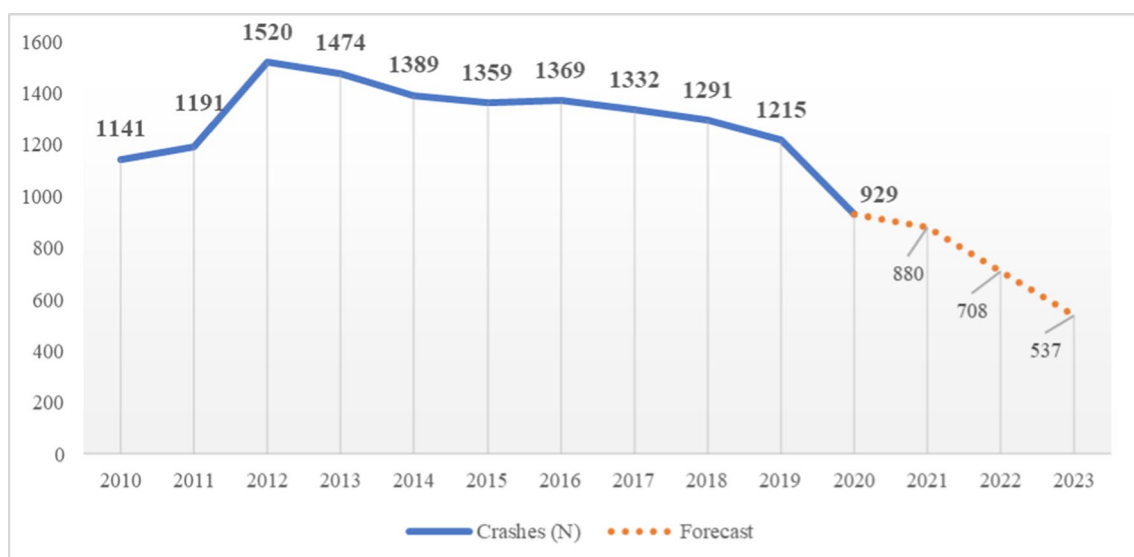


Fig. 1 Graphical distribution of crashes along observed period with forecast until year 2023

2. What are the latent structures of the primary components that underlie the investigated problem—namely, the quantity of pedestrian-related crash events on Vienna's roadways?
3. Which predictors, and to what extent, are they predictive of an ongoing high frequency of pedestrian-involved crashes in Vienna's urban centre?
4. In Vienna's urban city centre, which predictors and to what extent of influence distinguish injured pedestrians from those who are not?

Multivariate techniques constitute an increasingly important area of statistics in all branches of science. Offering a more complex examination of data and relationship between variables as well as possibility to predict future outcomes, it is the most appropriate methodology to tackle this complex problem of road safety, as previously confirmed by many researchers [3–6].

In response to defined knowledge gap, secondary analysis is applied on original police data using sophisticated multivariate techniques; factor and regression analytical methods to define important indicators, factors and predictors that are significantly influencing increase of the size of the number of crash event involving the most vulnerable traffic participants—pedestrians.

Main goal of this research is pointing out the leading characterizations in crashes with the aim to answer what is trend in crash occurrence during 2010–2020 inside Vienna municipality. Is there significant and distinctive difference based on gender and age with specific conditions under which crashes are occurring influencing different injury degree. What is predicting whether pedestrian will be injured or not while participating in traffic in Vienna.

To address this main problem of research it is performed: inferential analysis with focus on determining statistically significant differences, if any, among pedestrians on gender, age and injury severity; first order factor analysis method of main components to get principal dimension of crashes on variables that significantly characterize the crash event; second order factor analysis to uncover underlying structures in crash event involving pedestrians, and regression analysis to explain causal interconnection of continuous high number of crashes with certain predictors. Discriminant analysis to determine which variables will predict injuries in pedestrians using mathematical formula that combines a set of predictor variables.

On the original data collected by the police after each crash involving pedestrians, which occurred at the intersections/roads of Vienna between 2010 and 2020, it's applied cleaning, rearranging and recoding on all variables, thus adapting them for IBM's SPSS program to apply the more complex data processing techniques that this program allows. IBM's statistical software SPSS version 28 is used for all data analysis. It is the most used software by scientist all over the world due to its unique features for quantitative in-depth data analysis. Models developed using SPSS are essential tool in quantitative research because they allow to capture patterns that would not be obvious otherwise, make predictions and test whether our theories about the world stand up to empirical scrutiny [7].

Inferential analysis is used to conclude (inferentia lat. = reasoning) whether the resulting difference between the two sets of entities is statistically significant or not. To check the significance of differences, chi square test is used and from its size and size of the number of entities, the coefficient contingent is calculated, which shows whether the modalities in one independent variable significantly associate (correlate) with the modalities of the dependent variable. Inferential analysis used in this research was to test whether there is statistically significant difference based on gender and age when crossed with other variables.

Carrying out a multivariate technique—first order factor analysis, is aimed to uncover the latent structure that reveals the essential main components on which the investigated phenomenon stands—the crashes involving pedestrians in Vienna and their continuum reflected in high numbers recorded every year.

The scores coming from first order factor analysis were saved and on them the second order factor analysis was carried out that showed the latent structures lying in direct cause of crashes involving pedestrians in Vienna.

Multivariate analytical procedure allows application of multiple regression analysis adapted for categorical variables. This dependent technique is used to explain the variability of the criterion variable by variability of predictor variables and is considered the most important part of scientific research.

According to Blaikie [8], every scientific research should seek to answer the question: why something appears, happens, exists, or evolves. That's why it's necessary to carry out explanatory analysis. Explanatory analysis as one that seeks explanation, gives an answer to the question of what causes the researched phenomenon. The statistical procedure by which this is achieved is named regression analysis. This analysis in its multiple form results in the identification of essential predictors that significantly influence the variation of the investigated phenomenon and predict its movement and growth.

In this research, criterion variable is constructed by taking year with the highest number of crashes as modality eleven and the year with the lowest number as modality one while all other years are distributed among them in terms of growth by size expressed in number. This approach allowed getting variable that vary in number of crashes to show what significantly affected the greater growth of crashes in Vienna involving pedestrians.

As predictor variable in this multiple regression analysis using IBM'S CATREG are used predictors that predict variations of the mentioned criterion variable. The criteria behind what predictors to use is based on predictors obtained by variance analysis (F-test significant at the probability level of less than 5%). To look at the magnitude of the influence of individual significant predictors it's applied the stepwise procedure. With this procedure, some predictors are gradually included in the regression equation in the order of significance (determined by the size of the F-test) and their multiple correlation (R) is calculated to expresses the determination coefficient ($R^2 \times 100$) which tells us what the percentage of the common variance of the criterion variable and the individual group of predictors. In this way, with this stepwise procedure, we get an insight into how much individual predictors contribute to the explanation of the variation of the criterion variable by their inclusion in the regression equation. Through the height of the coefficient of the multiple correlation and the coefficient of determination, it's looked at what contributes to the explanation of the continuous high number of traffic crashes involving pedestrians in recent eleven years of observation.

Reasons for using discriminant analysis is building a predictive model for group membership as this model is composed of discriminant factors/function where it's linear combinations of the predictor variables provide best discrimination between the groups. In discriminant analysis we looked to see how we can best separate (or discriminate) a set of groups using several predictors. In some senses we're doing the reverse of the MANOVA: in MANOVA we predicted a set of outcome measures from a grouping variable, whereas in discriminant analysis we predict a grouping variable from a set of outcome measures [9].

The motivation to engage in this research stemmed from the fact that the continuous high and total number of crashes did not receive the necessary social attention and that to the authors best knowledge, there are no recent studies that have more seriously explored this phenomenon.

3 Inferential analysis results

Pedestrians gender showed statistically significant differences when crossed with variables referring to weather conditions, crash type, street type, intersection regulation, traffic lights and "hit and run" crash type, to mention few, but strength of association expressed through coefficient of contingency was weak ($CC < 0.1$). These results do not allow generalization but only shows tendency that female pedestrians are involved in crashes during early hours in foggy days while male pedestrians are with tendency to be involved in crashes during later hours in a day and sunny weather.

When it comes to pedestrians age, there is tendency that age differ significantly when crossed with temporal variables (weekend, seasons, weather), personal characteristics (gender, alcohol involved, alcohol test refused, disobeying traffic rules) and physical (intersection regulation, street type, road layout, conditions of road, traffic lights, traffic zone, road path). Due to $CC < 0.1$ we can talk about tendency and not firmly make conclusions. Younger pedestrians are with tendency to be involved in crashes during spring when drivers missed to indicate direction of travel change while older during winter when driver disregarded red traffic light.

Medium strength of association at level of contingency coefficient ($0.2 \leq CC < 0.3$) allows conclusions that crashes involving older pedestrians will be during weekends in later months of the year in town zones with higher speed allowed and highest frequency of crashes during dark conditions with artificial lighting. Older pedestrians at statistically significant level are more involved in crashes at roundabouts and crossings with offset branches, predominately at one-way streets, while younger are more involved in crashes at intersections. Police assumption on main cause of crash is predominately more correct when older pedestrians are involved in crash than younger. Crashes involving older pedestrians are including the most severe impairments, drugs, and alcohol while younger pedestrians' crashes are characterized with less severe impairment, health and medications.

Three variables referring to hours in a day are at statistically significant level differentiating young and old pedestrians involved in crashes in Vienna. Due to high contingency coefficient those associations are strong; each hour in a day $-\chi^2 = 3957,208$; $df(2254)$; $CC = 0.467$; $p = 0.001$; day split by four hours in a day $-\chi^2 = 1688,416$; $df(490)$; $CC = 0.326$; $p = 0.001$; day split by three hours in a day $-\chi^2 = 23,076$; $df(686)$; $CC = 0.352$; $p = 0.001$; irregular behavior $-\chi^2 = 1893,704$; $df(1176)$; $CC = 0.343$; $p = 0.001$; crash involving at least one person "on the way to school" $-\chi^2 = 4100,087$; $df(98)$; $CC = 0.473$; $p = 0.000$ and crash involving pedestrian younger than 14 years old wearing sign "on the way to school" $-\chi^2 = 4204,883$;

Table 1 Statistically significant chi squares, degrees of freedom, p-values ($p < 0.05$) and contingency coefficient in pedestrians' age groups crossed with "hit and run" crash (1.1), children involved (1.2), nationality (1.3) and injury degree (1.4)

1.1	No		Yes	
	f	%	f	%
Age in years				
≥ 10	1295	9.3	11	3.7
≥ 18	1679	12.1	15	5.1
≥ 25	1797	12.9	18	6.1
≥ 45	3576	25.8	36	12
≥ 65	3195	23	28	9.3
≥ 85	1995	14.3	6	1.9
86 ≤	308	2.2	1	0.4
nd	66	0.5	184	61.5
TOTAL	13911	100	299	100

 $\chi^2 = 0.6350,977$; $df(98)$; $CC = 0.556$; $p = 0.000$

1.2	No		Yes	
	f	%	f	%
Age in years				
≥ 10	0	0	1306	46.6
≥ 18	543	4.8	1151	41.2
≥ 25	1772	16	43	1.5
≥ 45	3403	30	209	7.5
≥ 65	3167	28	56	2.1
≥ 85	1978	17	23	0.8
86 ≤	309	2.7	0	0
nd	241	2.1	9	0.3
TOTAL	11413	100	2797	100

 $\chi^2 = 12101,53$; $df(98)$; $CC = 0.678$; $p = 0.000$

1.3	f	%	f	%	f	%
	Age in years					
≥ 10	103	13.4	772	7.9	431	12
≥ 18	91	11.8	1190	12.2	413	11
≥ 25	75	9.7	1215	12.5	525	14
≥ 45	127	16.5	2173	22.3	1312	35
≥ 65	122	15.8	2292	23.5	809	22
≥ 85	49	6.4	1743	17.9	209	5.7
86 ≤	6	0.7	291	3	12	0.3
nd	198	25.7	44	0.5	8	0.2
TOTAL	771	100	9720	100	3719	100

 $\chi^2 = 11,700,45$; $df(9604)$; $CC = 0.672$; $p = 0.001$

1.4	None		Light		Severe		Death	
	f	%	f	%	f	%	f	%
Age in years								
≥ 10	26	3.7	1090	10.1	187	7.3	3	2.5
≥ 18	39	5.4	1398	12.9	254	10	3	2.5
≥ 25	83	11.6	1486	13.8	240	9.4	6	5.1
≥ 45	243	34.1	2814	26	543	21.3	12	10.1
≥ 65	118	16.6	2446	22.6	630	24.7	29	24.3
≥ 85	19	2.7	1345	12.4	585	23	52	43.7
86 ≤	0	0	189	1.8	106	4.1	14	11.7

Table 1 (continued)

1.4	None		Light		Severe		Death	
	f	%	f	%	f	%	f	%
nd	186	26.1	61	0.6	3	0.1	0	0
TOTAL	714	100	10,829	100	2548	100	119	100

$\chi^2 = 3657,938$; $df(294)$; $CC = 0.452$; $p = 0.000$

$df(98)$; $CC = 0.478$; $p = 0.000$. Other influential variables that are at statistically significant level differentiate younger and older pedestrians involved in crashes are listed according to the size of chi-square test and presented with Table 1. Where for presentation purposes age groups are combined. Older pedestrians are predominately involved in crashes in later hours of the day while younger are involved in crashes in earlier hours of the day. Older pedestrians are severely injured in crashes involving distraction, alcohol/drugs, priority violation and pedestrian misconduct. Younger are with statistically significantly lighter injuries in crashes that involve irregular behaviour seen through modality technical failure of vehicle, obstacle and health impairment. Thus, crashes that result in more severe injuries involve foreign older pedestrians and child with mark "on the way to school" when driver left crash spot. Younger domestic pedestrians will suffer lighter injuries in crashes where drivers do not leave spot and do not include children. Undoubtedly those results give right to make this generalization to all Austrian population.

4 Factor analysis results

4.1 First order factor analysis

Carrying out a multivariate technique—factor analysis, is aimed to uncover the latent structure that reveals the essential main components on which the investigated phenomenon stands—the crashes involving the most vulnerable road users—pedestrians in Vienna and their continuum reflected in high numbers recorded every year. Table 2 presents eight factors structured from thirty-three variables included in first order factorization analysis of main components. Those eight factors explained half, 49% of total variance from all variables. Having categorical variable, author finds those findings beneficial and satisfactory.

The first main component is defined with variables describing person and crash, more specifically, travel movement in moment of crash and circumstance, lane, speed limit and road layout when specific disobeying traffic rules was registered. Mark variable "travel movement direction" is coded with seven modalities spanning from sharp left/right with modality two to standing still with modality seven. Because of mark variable composing this factor it is named "travel direction". This main component explains 11% of total variance. Higher crash occurrence involving pedestrians firstly relate to standing still at descending or ascending lane stripes where higher speed is allowed when distraction and/or irregular behaviour is recorded, prohibited use of phone while driving, disregarding traffic light and priority or having too small safety distance.

Second main component in first order is named "day vs night conditions" because mark variable "light conditions" describes amount of daylight recorded in time of crash. Variable is coded in four modalities where one is not defined, two is day, three is dusk and four is dark. Content of manifest variables are describing amount of lighting and hour of the day. Three manifest variables include description of alcohol impairment; was it involved and if testing was refused. Darker parts of the day in later hours show unsafe environment for pedestrians with higher crash occurrence involving alcohol and/or drugs impairment when alco-test was refused. This main factor explains 7% of total variance.

Third main component explains as 7% of total variance as previous one, meaning that they contribute with same percentage in explanation of high crash occurrence but this one is composed of manifest variables describing behaviour and road. Later years in observed period at intersections with none or manual regulation when police officer assumption was correct in defining main cause of crash was involving extremely risky behaviour; distraction, alcohol/drugs, priority violation and pedestrian misconduct in crashes at Vienna roads. Due to variables saturating this factor it can be called "observed years in dataset" as mark variable covers eleven years covered with this research where modality one is for Year 2010 and modality eleven is for Year 2020.

Table 2 Eight factor loadings greater than or equal to cut-off value 0.30 and percentage of variance explained coming from first order categorical principal component analysis (CATPCA)

Coefficient correlation with factor	Name of manifest variable	% of variance explained
<i>1st main component; FACTOR 1</i>		
0.921	Travel direction	11%
–0.914	Travel movement direction	
0.911	Lane	
0.846	Disobeying traffic rules	
0.638	Crash circumstance	
0.551	Road layout	
<i>2nd main component; FACTOR 2</i>		
–0.690	Speed limit	7%
0.606	Day vs night conditions	
0.580	Light conditions	
0.540	Time of day–split by 4 h	
0.523	Time of day–each hour	
0.485	Alcohol	
<i>3rd main component; FACTOR 3</i>		
0.783	Alco– test refused	7%
0.718	Observed years in dataset	
0.572	Year	
–0.548	Presumed main cause	
–0.406	Behaviour	
	Intersection shape	
<i>4th main component; FACTOR 4</i>		
0.490	Intersection regulation	6%
0.389	Crossing point regulation	
–0.381	Zebra crossing	
–0.376	Traffic lights	
–0.331	Alcohol presumed	
	Road conditions	
<i>5th main component; FACTOR 5</i>		
0.610	Weather	5%
0.610	Children	
0.338	Person on the way to school	
<i>6th main component; FACTOR 6</i>		
–0.390	Crash on the way to school	5%
–0.319	Child involved	
	Seasons	
<i>7th main component; FACTOR 7</i>		
0.585	Months	4%
0.561	Months + crashes count	
–0.449	“Hit and run crashes”	
–0.407	“Hit and run”	
	Age	
	Nationality	
<i>8th main component; FACTOR 8</i>		
0.652	Injury degree	4%
0.651	Standing still vs moving	
	Travelling “from”	
	Travelling “to”	
TOTAL = 49%		

Fourth main component is named “crossing point regulation” because manifest variables saturating this factor are describing if there was zebra crossing, condition of traffic regulation and police assumption of alcohol involvement. Very unsafe for pedestrian are points with zebra crossings when traffic regulation was working properly and police assumption on alcohol involvement was correct. Adding fog, snow and/rainy conditions that are creating slippery roads is environment for highest crash occurrence involving pedestrians. This main component explains 6% of total

variance and mark variable among which other manifest variables are saturating is “zebra crossing” that in its modality one is not defined, two–side way and three–zebra marking.

Fifth main component explains 5% of total variance and includes three different manifest variables containing child presence in highest crash occurrence. This factor can be named on mark variable “children” as it describes whether person involved in crash was with or without marking “on the way to school”.

Sixth main component contributes with same percentage to explanation of total variance, 5% respectively. This factor is specificatory due to two different modalities of manifest variables referring to months when crash occurrence involving pedestrian is higher, in earlier part of the year, respectively. This factor can be named “seasons” of the year as mark variable covers twelve months in the year where modality one is January and twelve is December.

Seventh main component explains 4% of total variance and is saturating manifest variables describing human factor and can be named “hit and run crashes” because mark variable describes whether crash involved or not driver that left crash spot. Highest crash occurrence involves older domestic pedestrians sustaining light injuries and drivers leaving crash spot. Crash involving drivers that are not leaving crash spot unfortunately consider as well severely injured young and foreign pedestrians.

Eight main component contribute as well with 4% in explaining of variation in total variance. Manifest variables are specificatory ones as they indicate West and North-West travelling directions as riskier for pedestrians among eight cardinal directions. This factor can be named “standing still vs moving” as first modality in both variables, “from” and “to” respectively, is standing and modality ten is North-West.

Those findings allow conclusion that pedestrians movement along Vienna is very unsafe imposing higher crash occurrence with exogenous variables describing weather conditions, seasons and time of the year along observed period with no natural light. Endogenous variables among others, warns of alcohol use while participating in traffic with violation of traffic rules. Nothing less important is finding on children involvement with higher crash occurrence. Police presence and right judgment confirms importance of continuous improvement in a way that crash data is collected and processed. Corrective measures directed to those main factors are hiding potential to half number of crashes involving pedestrians.

4.2 Second order factor analysis

Relatively rare are factorization approaches in research that go deeper in latent structures of highest crash occurrence involving pedestrians in Vienna when second order factors are defined. It is possible to save scores and on them perform another factor analysis in second order. This approach gives second order factors that are saturated with first order factors that are presented. The scores of those eight first order factor analysis was saved, and second order factor analysis is carried out. Results presented with Fig. 3 show four factors of second order factor analysis Table 3.

First main component in second order is in high correlation with fifth, second and seventh factor in first order factor analysis and can be named on mark variable “children”. Highest correlation is with mark variable due to coefficient

Table 3 Four factor loadings greater than or equal to cut-off value 0.50 and percentage of variance explained coming from second order categorical principal component analysis (CATPCA) on scores coming from first order

Coefficient correlation with second order factor	Name of first order factor	% of variance explained
<i>1st main component; FACTOR 1 s order</i>	<i>Children</i>	<i>29%</i>
0.921	Children–Factor 5 in first order	
–0.914	Day vs night conditions–Factor 2 in first order	
0.911	“Hit and run” crashes–Factor 7 in first order	
<i>2nd main component; FACTOR 2 s order</i>	<i>Seasons</i>	<i>17%</i>
–0.690	Seasons–Factor 6 in first order	
0.606	Standing still vs moving–Factor 8 in first order	
<i>3rd main component; FACTOR 3 s order</i>	<i>Observed years in dataset</i>	<i>16%</i>
0.783	Observed years in dataset–Factor 3 in first order	
0.718	Travel direction–Factor 1 in first order	
<i>4th main component; FACTOR 4 s order</i>	<i>Crossing point regulation</i>	<i>13%</i>
0.490	Crossing point regulation–Factor 4 in first order	
		TOTAL = 76%

$R=0.921$ and lighting conditions $R=-0.914$. These results show connection between highest crash occurrence involving children during day conditions when drivers left crash spot. Second main component is saturating factors from first order named seasons and travel and can be named “seasons” due highest correlation with factor six in first order factor analysis. This factor excludes human factor but show highest crash occurrence for pedestrians in later months of the year when travel direction was defined as standing still and/or travelling from/to North-West. Third main component in second order analysis is saturating factors describing observed eleven years observed period and can be named “observed years in dataset”. Highest correlation with factor is with third main component in first order analysis due to coefficient $R=0.783$ and with first component in first order, $R=0.718$. This shows that in latent structures of highest pedestrian-related crash occurrence are earlier years and moving straight ahead. Fourth main component is as well fourth factor in first order pointing out crossing point specifications as well in deep latent structures of highest crash occurrence involving pedestrians. This component can be named “crossing point regulation” after mark variable that combines “zebra crossing” indicating whether it was zebra marking or not, and traffic lights. This main factor explains 13% of common variance and reveals deepest latent structures in pedestrian-related crashes.

Generally, those findings allow conclusion that first factor is composed of relatively complex combination of three factors from first order analysis while the next two are saturating two factors each. Fourth component in second order analysis is in a way specification because it combines one factor from first analysis. For further comment it is important to highlight that in deeper factor layers we find those structures that explain the high number of crashes involving pedestrians in Vienna through manifest variables.

5 Regression analysis results

A total of thirty-three variables coming from previous factor analysis were included in regression analysis and F-test extracted those seventeen predictors using ANOVA procedure which significantly predict the increasing size of the crash event involving pedestrians. Results are presented with Table 4.

The first five predictors in the model; time of the day, travelling direction, seasons of the year, age and involvement of children, are keeping low impact on the size of the crash event explaining below 2% of the common variance at criterion variable. When intersection shape predictor is added in the model it is significant “jump” in common variance explained, 8,35% respectively. Roundabout and crossing with offset branches are more frequent crash spots involving pedestrians than intersections. Next predictors in model are increasing percentage of common variance explained. They are crash spot pathway; signalization, pathway, cycling path and crash spot surroundings. Physical specification is limiting area of action, but human factor extends those possibilities. Adding human behaviour in the model explains 35% of common variance

Table 4 Results of stepwise regression analysis in contribution of multiple coefficients to the criterion variable—the size of the number of crash event

Criterion variable	Predictor variables	F test	R	$R^2 \times 100$
Size of the number of crash events	1 Time of day—split by 4 h	F = 3.791	$r=0.012$	$R^2=0,0\%$
	2 Travelling “to”	F = 3.988	$R=0.115$	$R^2=1,32\%$
	3 Pre& post seasons	F = 4.014	$R=0.118$	$R^2=1,39\%$
	4 Age in years	F = 4.930	$R=0.118$	$R^2=1,40\%$
	5 Child involved	F = 6.876	$R=0.123$	$R^2=1,51\%$
	6 Intersection shape	F = 8.434	$R=0.289$	$R^2=8,35\%$
	7 Traffic lights	F = 10.021	$R=0.292$	$R^2=8,52\%$
	8 Pathway	F = 11.700	$R=0.315$	$R^2=9,92\%$
	9 Cycling	F = 15.698	$R=0.345$	$R^2=11,90\%$
	10 Road space	F = 16.682	$R=0.362$	$R^2=13,10\%$
	11 Behaviour	F = 21.796	$R=0.590$	$R^2=34,81\%$
	12 Presumed alcohol	F = 26.602	$R=0.594$	$R^2=35,28\%$
	13 Intersection regulation	F = 35.769	$R=0.601$	$R^2=36,12\%$
	14 Main cause presumed	F = 53.118	$R=0.743$	$R^2=55,20\%$
	15 Lighting conditions	F = 77.732	$R=0.761$	$R^2=57,91\%$
	16 Lane	F = 530.549	$R=0.919$	$R^2=84,45\%$
	17 Travel movement direction	F = 2904.051	$R=0.936$	$R^2=87,60\%$

causing the second “jump”. Thus, higher contribution in increasing crash occurrence is related to distraction, alcohol, drugs, violation of priority and pedestrian misconduct. Alcohol presumed and intersection regulation are next predictors in the model. Presuming main cause of crash is defined by police officer at the spot and causes “jump” in increasing % of common variance explained, from 36 to 55% respectively. Importance and integration of police observations, not rarely and experience, proves to be important predictor in the size of crash event. Lighting conditions is next predictor in model explaining 58% of common variance, but when lane where crash occurred is added in the model, we have final big “jump” explaining 84% of common variance. Least predictor in model is general movement direction reaching total of 87,60% of common variance explained—the size of the crash event involving pedestrians in Vienna.

Observing all predictors in model that are important for predicting the size of crash event involving pedestrians, two are human characteristics (age and behaviour), seven are road specifications and three precipitations. Highest jumps in the model are when intersection shape, behaviour and police judgement is included in the model. This clearly indicates interconnection of environment and traffic participants. If we want to eliminate more than 87% of crashes involving pedestrians in Vienna, corrective measure directed towards those predictors are defined starting point.

6 Discriminant analysis results

On pedestrian dataset, chi-square was calculated on all variables. Application of discriminant analysis on those significant variables enabled definition of the ones that at statistically significant level define whether pedestrian will sustain injury or not. Further application of stepwise allowed identification of smallest number of predictors in the model. In Table 5 it is presented twelve predictors included in stepwise discriminant function that explained 100% of variance while canonical $R^2 = 0.357$ suggest that model explains 35% injured and not injured pedestrians as variation in grouping variable. Discriminant function for injured and not injured pedestrians is significantly differentiated, Wilk's Lambda = 0.873, $\chi^2(12) = 1932.188$, $p = 0.000$.

Law prescribes obligation of not leaving crash spot and this behaviour is taught in driving schools. With the added involvement of children, we can picture very dangerous traffic situation. Very systematic and well thought system from passing driving exam and getting driving license if probation period was without serious offense shows need for something similar in older age. Variable referring to crash circumstance combines irregular behaviour (of pedestrians until year 2017), visual restrictions, boarding or disembarking from public transportation and distraction. In town speed limits are generally with maximum 50 km/h but in shared space go up to maximum of 30 km/h. This shows that speed with other factors play a discriminant role in outcome for pedestrians seen as whether they will end up injured or not. Next is intersection shape and registered disobeying of traffic rules. Seasons of the year, gender and impairment are followed with road conditions and crash type that when combined all together, give very vivid picture discriminating against injured pedestrians and the ones that are not.

Standardized canonical discriminant function coefficients provide an index of importance of each predictor in the model. Speed limit, seasons of the year, intersection shape and impairment are contributing to variate in a negative way while all other predictors are contributing to variate in a positive way. Positive predictors “hit and run” crash when child

Table 5 Variables included in stepwise discriminant analysis with statistically significant F and p-values ($p < 0.05$) in pedestrians' group given in test of equity in group of means

Variables in analysis	F test	Sig
1 “Hit and run”	F = 1605.932	0.000
2 Child involved	F = 41.316	0.001
3 Age	F = 206.670	0.001
4 Crash circumstance	F = 51.590	0.001
5 Speed limit	F = 15.741	0.001
6 Intersection shape	F = 42.255	0.001
7 Disobeying traffic rules	F = 33.740	0.001
8 Seasons	F = 36.006	0.001
9 Gender	F = 35.132	0.001
10 Impairment	F = 14.578	0.001
11 Road conditions	F = 32.516	0.001
12 Crash type	F = 12.773	0.001

was involved in combination with age are making the largest contribution whether pedestrian will be injured or not. Despite lower but still significant predictors for injuries in pedestrians are as well crash circumstance, multiple levels of disregarding traffic rules, road conditions, gender and crash type.

Values in structure matrix as the canonical variate correlation coefficients indicate substantive nature of variates. Considering not proscribed rule that each variable with value of correlation coefficient ≤ 0.30 is considered significant, Pearson's coefficients are structuring discriminant difference that in its respect to aforementioned are "hit and run" and age expressed in years, both with positive effect.

Further interpretation of discriminant analysis results defines pedestrians that are injured against the ones that are not in sense of profiles using group means of predictor variable. In pedestrian sample injured produce a mean of 1.660 while not injured produce a mean of -0.088 and this allows to make conclusion how injured pedestrians are more discriminatory than not injured. Classification function coefficients show higher negative sign for crash circumstances and disobeying traffic rules for not injured pedestrians, and they are the only two predictors with negative sign. Predictors with positive sign and expressed with higher values for not injured pedestrians are as well speed limit, intersection shape, impairment and season in the year. Gender, age, road conditions, crash type, "hit and run" crashes when child was involved are expressed with positive sign and injured pedestrians.

7 Discussion

Analysing large sets of data is complex and time-consuming process. Statistical package for the social sciences (SPSS) combines software programs in single place providing different techniques to analyse, transform and produce patterns between variables [4, 10–14] and it is widely used among researchers to tackle road safety which makes it reasonable and valid selection for this research.

Different approaches and varying trends can be identified in road safety literature spanning from exploration of various aspects of crashes and fatalities [15, 16], driver behaviour and traffic violations [17], analysis of factors affecting road safety [18] to policy awareness of traffic regulations and factors that can improve safety [19].

Statistically significant differences, chi-square, proved age, gender, road familiarity, risky behaviour and driving habits contributing to crash occurrence [20] showing similarities in results with Vienna where it is proved older pedestrians are severely injured in crashes involving distraction, alcohol/drugs, priority violation and pedestrian misconduct while younger pedestrians predominately obtain light injuries in crashes that involve irregular behaviour in urban city centre of Vienna. Crashes in urban environments are prevailing in business zones and highly populated commercial housing zones [16]. Despite low maximum speed allowed in urban area, generally maximum of 50 km/h, speed is still risk factor in crashes, increasing likelihood of crash and severity of injury. According to Pan American Health Organization [21], if a child crosses the road 13 m in front a car travelling at 30 km/h, it can stop just before hitting the child. However, if the car's speed is 50 km/h or more, the child will be hit by the car and there will be little chance of survival. Speed is recognized and implemented as one of pillars in Sweden well known Safe System approach where only zero is acceptable number when counting deaths and injuries in road network. It also acknowledges shared responsibility and human nature to make mistakes. Analysis found that the risk of fatality reaches 50% at estimated impact speed of 59 km/h making more than reasonable implementation of lower speed limits, maximum of 40 km/h in pedestrian active area [22].

Multivariate statistical approach of principal component analysis using fourteen explanatory variables to analyse crash occurrences on interstate highways revealed the facts that the two human factors are linearly combined to form a single factor, and the prevailing weather condition is correlated with the pavement surface condition [23]. Among eight components in first order factor analysis only two factors excluded human factor, sixth (seasons of the year) and eight (travel directions), respectively. Going to deeper layers with second order factor analysis among four factors, was especially worrying seeing first main component saturating manifest variables referring to whether children was involved and/or on the way to school. Those results are supported with findings where sixteen out of 30 children (53.3%) were injured on their way to school, 3 (10.0%) on their way to kindergarten, nine out of 30 children (30%) were injured when crossing the street from behind the bus or the tram, five children (16.7%) were injured by buses, 6 (20.0%) by trams and 19 (63.3%) by passing motorized vehicles and twelve (40.0%) were injured in the presence of accompanying adults [24]. Multivariate analysis revealed that vehicle travel speed greater than 30 km/h, pedestrian age less than 5 years, time of day either early morning or late afternoon, residential zone, type of road including collectors and major roads, and centre travel lanes were associated with greater severity of injury [25]. Results showed that risky behaviours while travelling

as a pedestrian were more widespread among ninth grade students and in student population it was found greater tendency to justify risky behaviour [26].

Stepwise regression showed significant “jumps” in total variance explanation causing predictor variables on intersection shape, irregular behaviour, presumed main cause of crash by police officer and lane. The importance of age-related effects in injury severity is verified by hierarchical and principal components logistic regression models, amplifying findings of exploratory stepwise logistic analysis where variations in findings resulted when the population was divided by gender [27]. The study using Louisiana pedestrian fatal and injury crash data crashes at high-speed segments are associated with no physical separation, dark-no-streetlight conditions, 3-leg intersection geometry and combination of driver irregular behaviour and pedestrian intoxication [28]. Three and four-leg intersections unsignalized intersections differ in number of conflict points, nine vs thirty-two respectively (including crossing, merging and diverging). Those intersection designs are widely implemented in urban areas. Observing only crossing conflict points there are three at 3-leg intersections and sixteen at 4-leg. Cause of crashes differ among them as showed that crashes at 3-leg intersections as main cause have minor street passing through intersection while at 4-leg it is passing of the street [29]. When participating in traffic more often than not, pedestrians are crossing street in prohibited way. If the pedestrian is outside the crosswalk area with no lighting conditions at dark it is associated with high number of crashes [30]. Several important factors affecting crash severity at unsignalized intersections include traffic volume on major approach, number of through lanes on minor approach, left and right shoulder width, number of right and left turn lanes on major approach where young drivers were as well associated with less fatal probability when compared to other age groups [31]. Using electronic gadgets, not using pedestrian crosswalk and using a mobile phone was reported by participants when examining road risk behaviour [32].

Having analysed the irregular movements according to the type of crossing it was found that car drivers behave the most irregularly at crossings without a refuge island [33]. Maintaining high speed (even exceeding the speed limit of 50 km/h) is the signal from the drivers that they do not intend to give way to the pedestrian at the zebra crossing and the conclusion is that encounters between cars and pedestrians at the zebra crossing are critical situations in which the driver has to be influenced before he reaches the decision zone at 50 to 40 m before the zebra crossing in order to prevent the ‘signalling by speed’ behaviour [34].

The severity of the injured pedestrians is related to many factors, including vehicle speed, the angle of impact of the vehicle upon the pedestrian, the centre of gravity of the pedestrian, the part of the body that first comes into contact with the vehicle, the part of the vehicle the pedestrian impacts first, and the vehicle design [35]. Discriminatory analysis showed factors defining injury in pedestrians at Vienna city centre. Hit and run drivers appear to run more often when they are at fault and intoxicated and when it is likely they can escape detection, at night [36]. Majority of crashes occurred at roads (83%) between 12 p.m. and 6 p.m. (53%) involved pediatric pedestrian injuries in primary-schoolers walking home from school unaccompanied by adults [37]. Vulnerability and risk exposure showed that children are more likely to be injured during weekends in residential zones and males during afternoon in arterial streets [38]. Similar results are coming from Finland study where seriously injured pedestrians occur more often than fatalities with passenger cars, in wet road conditions, with streetlights lit, in temperatures from -3 to 3 °C, in speed limit zones of 40–70 km/h, on pedestrian crossings, and in inner urban areas [39]. Drunk driving was a significant risk factor for fatal injuries among vulnerable road users [40] but intoxicated pedestrians are significantly more likely to sustain fractures, require operative procedure or intensive care unit admission [41]. On one side more loudly, countries are promoting walking and its undoubted health benefit but among fatally injured pedestrians tested for alcohol and drugs, 39.7% and 43.4% tested positive [42] shows attention direction for prevention and mitigation of this safety problem. Reversing is just one of driving behaviour captured to influence crashes involving pedestrians in Vienna and according to Decker et al. [43] majority of crashes occur at low impact speed during day where pedestrians are in greater risk of severe injury at any body region. Unsafe lane changes included distraction due to mobile phone [44] and probability that driver will encounter a rare event during driving increases faster than kilometres driven [45]. Red light runners, as a group, were younger, less likely to wear seat belts, had poorer driving records, and drove smaller and older vehicles than the drivers who did not run red lights [46].

In European Road Safety Observatory [47] during year 2021 pedestrian fatalities made up 1/5 of all road fatalities. Younger pedestrians are more likely to commit a violation in the same situation than older ones, although their physical condition and such as better mobility lead them less to be involved in conflicts and crashes [48]. From the perspective of driving groups, young drivers, notice drivers, and older drivers are the main groups to be intervened, mainly with education and training [49]. Road safety education and physical environment must be considered together, in a new way that will give desired results. Pre-and post-license education by people of all ages led to improvements in secondary outcomes (e.g., self-perceived driving abilities, behind-the-wheel driving performance, and even a small decrease

in traffic offenses) but education was not effective in reducing crashes or injuries, either at the individual or community level [50–52]. Evidence suggest effectiveness of road safety education interventions at best but in short term [50].

Training programs are out of scope of this research article but addressing critical age and experience related factors as proved with this research might be good direction in redefining our understanding of safety problem. Another problem not considered but emerging and passing under noticed are significant injuries inflicted on passengers of public buses [53].

Urban environment needs more attention and studies if we as society, independent of place where we live want safe road networks were reaching zero deaths and injuries is target for all stakeholders.

8 Conclusion and recommendations

In relation to previously set goal and research problem, those results allow following conclusions that can be seen as answers to previously defined research questions.

Inferential analysis showed there is significant and distinctive difference based on gender and age with specific conditions under which crashes are occurring and influencing different injury degree. Crashes that result in more severe injuries involve foreign older pedestrians and child with mark “on the way to school” when driver left crash spot, while the opposite description is for light injuries.

First order factor analysis revealed eight factors as latent structures of main components on which the researched problem lies—the size of the number of the crash event at Vienna’s roads involving pedestrians. Second order factor analysis showed four factors lying in depth latent structures of crashes involving pedestrians and they are saturating on personal characteristics (impairment; children and person), precipitation (seasons, hour and travel), and physical characteristics (crossing point).

Observing all predictors in stepwise regression model that are of importance for predicting the size of crash event involving pedestrians, two are human characteristics (age and behaviour), seven are road specifications and three precipitations. Highest increase in percentage of variance explained is causing children involvement, person behaviour, police assumption on main cause of crash and lane.

Predictors with positive sign that at statistically significant level are predicting pedestrian’s injuries are physical characteristics (gender, age and child involvement) that under specific road conditions end in different crash type, including “hit and run” crash.

Multivariate techniques used in this research for uncovering new knowledge of contributing factors in pedestrian’s group has been useful for all analysis performed because they revealed underlying risk differences and influences.

Methodology applied proved to be valuable for characterization of circumstances for crashes and pedestrian injuries in metropolitan area of Vienna and has potential to be used in other developed urban centres.

Undoubtedly this research contributed to explanation of underlying factors in crashes involving pedestrians, but limitations should be acknowledged. Further research is needed in other traffic groups providing more insights into traffic situation. Furthermore, safety comparison of similar developed urban city centres like Vienna could provide more conclusions and extent knowledge in understanding this safety problem. Facing new travel modes, distinctions on “new age” soft travelling modes is necessary for obtaining general safety image. Existing measures and law regulations are playing a role in safety culture and are indicators of very well thought and structured National strategy but are out of the scope of this research.

Acknowledgements I as author confirm that there are no known conflicts of interest associated with this publication and there has been no financial support for this work that could have influenced its outcome, nor has been published elsewhere. First I would like to thank *BUNDESANSTALT STATISTIK ÖSTERREICH, Direktion Raumwirtschaft, Kraftfahrzeuge, Straßenverkehrssicherheit, Guglgasse 13, 1110 Wien*. Thank you for support you provided. I am also grateful for insightful comments offered by the anonymous peer reviewers. Your valuable insights have contributed to better presentation of my work. Next I would like to express gratitude to my professor Obradovic and Kristina. Biggest thanks to my parents and brother for supporting me whatever I pursue.

Authors’ contributions RM data acquisition, data validation, research concept and methodology, proof reading.

Funding Not applicable, no funding.

Data availability Due to legal restrictions, supporting original police data is not available to be shared publicly in any form.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Author, Radmila Magusic gives consent for publication.

Competing interests Author confirms there is none relevant financial or non-financial competing interest to report.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. International transport forum. austria: road safety country profile. 2021. <http://www.itf-oecd.org/austria-road-safety> Accessed 25 Jan 2023
2. Austrian road safety strategy 2021–2030. (nd.). <https://www.bmk.gv.at/en/topics/transport/roads/safety/vss2030.html> Accessed 25 Jan 2023
3. Everitt BS, Landau S. A handbook of Statistical analyses using SPSS. In Chapman and Hall/CRC eBooks. 2003. <https://doi.org/10.1201/9780203009765>.
4. Field AP. Discovering statistics using IBM SPSS Statistics: and sex and drugs and rock 'n' roll, 4th edition. In Sage eBooks. 2013. <http://sro.sussex.ac.uk/id/eprint/43894/>
5. Bala J. Contribution of SPSS in social sciences research. *Int J Adv Res Comput Sci*. 2016; 7(6), 250–254. <https://api.semanticscholar.org/CorpusID:150245130>
6. Rahman A, Muktadir M. SPSS: an imperative quantitative data analysis tool for social science research. *Int J Res Innovat Soc Sci*. 2021;5(10):300–2. <https://doi.org/10.47772/ijriss.2021.51012>.
7. Tolmie A, Muijs D, McAteer E. Quantitative methods in educational and social research using SPSS. In Open University Press eBooks (Issue 1). 2011. <https://api.semanticscholar.org/CorpusID:60713699> Accessed 6 Jan 2023
8. Blaikie N. Designing social research: The Logic of Anticipation. Polity. 2009
9. Field A. Discovering statistics using SPSS. SAGE Publications. 2009
10. Jolliffe I. Principal component analysis. Springer Science & Business Media. 2002
11. Acton C, Miller R, Maltby J, Fullerton D. SPSS for Social Scientists. Red Globe Press. 2009. <https://books.google.hr/books?id=ECBHEAAQBAJ>.
12. Hair JF. Multivariate data analysis. Pearson. 2010.
13. Rovai AP, Baker JD, Ponton MK. Social science Research design and statistics: a Practitioner's Guide to Research Methods and SPSS Analysis. Watertree Press LLC. 2013
14. Ho R. Handbook of Univariate and Multivariate Data Analysis with IBM SPSS. In Chapman and Hall/CRC eBooks. 2013. <https://doi.org/10.1201/b15605>
15. Alkaabi K. Identification of hotspot areas for traffic accidents and analyzing drivers' behaviors and road accidents. *Transportation Res Interdiscipl Perspect*. 2023;22: 100929. <https://doi.org/10.1016/j.trip.2023.100929>.
16. AlKetbi LMB, Grivna M, Dhaheri SA. Risky driving behaviour in Abu Dhabi, United Arab Emirates: a cross-sectional, survey-based study. *BMC Public Health*. 2020. <https://doi.org/10.1186/s12889-020-09389-8>.
17. Mokarami H, Alizadeh SS, Pordanjani TR, Varmazyar S. The relationship between organizational safety culture and unsafe behaviors, and accidents among public transport bus drivers using structural equation modeling. *Transport Res F Traffic Psychol Behav*. 2019;65:46–55. <https://doi.org/10.1016/j.trf.2019.07.008>.
18. Li H, Liu S. SPSS-Based Analysis of Factors Affecting Road Traffic Safety. 2013. Doi: <https://doi.org/10.1061/9780784413159.228>.
19. Bendak S. Seat belt utilization in Saudi Arabia and its impact on road accident injuries. *Accid Anal Prev*. 2005;37(2):367–71. <https://doi.org/10.1016/j.aap.2004.10.007>.
20. Bucsuházy K, Matuchová E, Zúvala R, Moravcová P, Kostíková M, Mikulec R. Human factors contributing to the road traffic accident occurrence. *Transportat Res Procedia*. 2020;45:555–61. <https://doi.org/10.1016/j.trpro.2020.03.057>.
21. Pan American Health Organization. (2018). Speed and road crashes. PAHO/NHM/18–007. Washington, DC: Pan American Health Organization.
22. Hussain Q, Feng H, Grzebieta R, Brijs T, Olivier J. The relationship between impact speed and the probability of pedestrian fatality during a vehicle-pedestrian crash: a systematic review and meta-analysis. *Accid Anal Prev*. 2019;129:241–9. <https://doi.org/10.1016/j.aap.2019.05.033>.
23. Kassu A, Hasan M. Identifying the principal factors influencing traffic safety on interstate highways. *SN Appl Sci*. 2019. <https://doi.org/10.1007/s42452-019-1796-2>.
24. Unger R, Eder C, Mayr J, Wernig J. Child pedestrian injuries at tram and bus stops. *Injury-Int J Care Injured*. 2002;33(6):485–8. [https://doi.org/10.1016/s0020-1383\(02\)00051-7](https://doi.org/10.1016/s0020-1383(02)00051-7).

25. Pitt R, Guyer B, Hsieh CC, Malek M. The severity of pedestrian injuries in children: an analysis of the Pedestrian Injury Causation Study. *Accid Anal Prev.* 1990;22(6):549–59. [https://doi.org/10.1016/0001-4575\(90\)90027-i](https://doi.org/10.1016/0001-4575(90)90027-i).
26. Salduccio A, Abati D, Bina M, Bertani B, Calarco S, Balzarotti S, Biassoni F. Young pedestrians' behaviours and risk perception: a pilot study with Italian early adolescents. *Transport Res F Traffic Psychol Behav.* 2022;90:512–23. <https://doi.org/10.1016/j.trf.2022.06.006>.
27. Mercier C, Shelley M, Rimkus JB, Mercier J. Age and gender as predictors of injury severity in head-on highway vehicular collisions. *Transp Res Rec.* 1997;1581(1):37–46. <https://doi.org/10.3141/1581-05>.
28. Hossain A, Sun X, Zafri NM, Codjoe J. Investigating pedestrian crash patterns at high-speed intersection and road segments: findings from the unsupervised learning algorithm. *Int J Transportat Sci Technol.* 2023. <https://doi.org/10.1016/j.ijst.2023.04.007>.
29. Stanton N. Advances in human aspects of transportation: Proceedings of the AHFE 2020 Virtual Conference on Human Aspects of Transportation, July 16–20, 2020, USA. Springer Nature. 2020.
30. Das S, Tamakloe R, Zubaidi H, Obaid I, Alnedawi A. Fatal pedestrian crashes at intersections: trend mining using association rules. *Accid Anal Prev.* 2021;160: 106306. <https://doi.org/10.1016/j.aap.2021.106306>.
31. Haleem K, Abdel-Aty M. Examining traffic crash injury severity at unsignalized intersections. *J Safety Res.* 2010;41(4):347–57. <https://doi.org/10.1016/j.jsr.2010.04.006>.
32. Narváez YV, Sierra VP, Cárdenas FP, Ramos LR, González BZ, Martínez JIV, Aranda ÓM. Road risk behaviors: pedestrian experiences. *Traffic Inj Prev.* 2019;20(3):303–7. <https://doi.org/10.1080/15389588.2019.1573318>.
33. Makó E, Szakonyi P. Evaluation of human behaviour at pedestrian crossings. *Transport Res Proced.* 2016;14:2121–8. <https://doi.org/10.1016/j.trpro.2016.05.227>.
34. Várhelyi A. Drivers' speed behaviour at a zebra crossing: a case study. *Accid Anal Prev.* 1998;30(6):731–43. [https://doi.org/10.1016/s0001-4575\(98\)00026-8](https://doi.org/10.1016/s0001-4575(98)00026-8).
35. Chakravarthy B, Lotfipour S, Vaca FE. Pedestrian injuries: emergency care considerations. *California J Emer Med.* 2007;8(1):15–21.
36. Solnick SJ, Hemenway D. The hit-and-run in fatal pedestrian accidents: victims, circumstances and drivers. *Accid Anal Prev.* 1995;27(5):643–9. [https://doi.org/10.1016/0001-4575\(95\)00012-0](https://doi.org/10.1016/0001-4575(95)00012-0).
37. Feng XYJ, Nah SA, Lee YT, Lin Y, Chiang LW. Pedestrian injuries in children: who is most at risk? *Singapore Med J.* 2015;56(11):618–21. <https://doi.org/10.11622/smedj.2015170>.
38. De Santiago Armenta-Ramirez I, Reyes-Castro PA, Zúñiga-Terán AA, Olmedo-Munoz M. The urban structure and pedestrian injuries: a typological analysis of pedestrian crashes in the city of Hermosillo. *Mexico Traffic Injury Prevent.* 2023;24(5):428–35. <https://doi.org/10.1080/15389588.2023.2204386>.
39. Malin F, Silla A, Mladenović MN. Prevalence and factors associated with pedestrian fatalities and serious injuries: case Finland. *Eur Transport Res Rev.* 2020. <https://doi.org/10.1186/s12544-020-00411-z>.
40. Lin H, Chan C, Wiratama BS, Chen P, Wang M, Chao C, Saleh W, Huang H, Pai CW. Evaluating the effect of drunk driving on fatal injuries among vulnerable road users in Taiwan: a population-based study. *BMC Public Health.* 2022. <https://doi.org/10.1186/s12889-022-14402-3>.
41. Tonellato DJ, Ransohoff JR, Nash CC, Melanson SE, Petrides AK, Tolan NV, Goldberg SA, Boyer EW, Chai PR, Erickson T. Traumatic pedestrian and bicyclist injuries associated with intoxication. *Am J Emerg Med.* 2021;45:192–5. <https://doi.org/10.1016/j.ajem.2020.08.024>.
42. Vanlaar W, Hing MM, Brown S, McAteer H, Crain J, McFaul SR. Fatal and serious injuries related to vulnerable road users in Canada. *J Safety Res.* 2016;58:67–77. <https://doi.org/10.1016/j.jsr.2016.07.001>.
43. Decker S, Otte D, Cruz D, Müller C, Omar MA, Krettek C, Brand S. Injury severity of pedestrians, bicyclists and motorcyclists resulting from crashes with reversing cars. *Accid Anal Prev.* 2016;94:46–51. <https://doi.org/10.1016/j.aap.2016.05.010>.
44. Shawkly M. Factors affecting lane change crashes. *IATSS Res.* 2020;44(2):155–61. <https://doi.org/10.1016/j.iatssr.2019.12.002>.
45. Elvik R. Driver mileage and accident involvement: a synthesis of evidence. *Accid Anal Prev.* 2023;179: 106899. <https://doi.org/10.1016/j.aap.2022.106899>.
46. Retting RA, Williams AF. Characteristics of red light violators: results of a field investigation. *J Safety Res.* 1996;27(1):9–15. [https://doi.org/10.1016/0022-4375\(95\)00026-7](https://doi.org/10.1016/0022-4375(95)00026-7).
47. European Commission. Facts and Figures Pedestrians. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. 2021.
48. Sheykhfarid A, Haghighi F, Papadimitriou E, Van Gelder P. Review and assessment of different perspectives of vehicle-pedestrian conflicts and crashes: passive and active analysis approaches. *J Traffic Transportat Eng.* 2021;8(5):681–702. <https://doi.org/10.1016/j.jtte.2021.08.001>.
49. Yang M, Bao Q, Shen Y, Qu Q. Thirty years of research on driving behavior active intervention: a bibliometric overview. *J Traffic Transportat Eng.* 2023. <https://doi.org/10.1016/j.jtte.2023.06.002>.
50. Korner-Bitensky N, Kua A, Von Zweck C, Van Benthem K. Older driver retraining: an updated systematic review of evidence of effectiveness. *J Safety Res.* 2009;40(2):105–11. <https://doi.org/10.1016/j.jsr.2009.02.002>.
51. Akbari M, Lankarani KB, Heydari ST, Motevalian SA, Tabrizi R, Sullman MJM. Is driver education contributing towards road safety? a systematic review of systematic reviews. *J Injury Viol Res.* 2021. <https://doi.org/10.5249/jivr.v13i1.1592>.
52. Poulter D, McKenna F. Evaluating the effectiveness of a road safety education intervention for pre-drivers: an application of the theory of planned behaviour. *Br J Educ Psychol.* 2010;80(2):163–81. <https://doi.org/10.1348/014466509x468421>.
53. Halpern P, Siebzeiner MI, Aladgem D, Sorkine P, Bechar R. Non-collision injuries in public buses: a national survey of a neglected problem. *Emerg Med J.* 2005;22(2):108–10. <https://doi.org/10.1136/emj.2003.013128>.