Drivers' attitudes to road safety in the South East Asian cities of Jakarta and Hanoi

Thibenda, Moris; Wedagama, Dewa Made Priyantha; Dissanayake, Dilum

DOI:
10.1016/j.ssci.2022.105869

Safety Science 155 (2022) 105869

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Drivers' attitudes to road safety in the South East Asian cities of Jakarta and Hanoi: Socio-economic and demographic characterisation by Multiple Correspondence Analysis

Moris Thibenda [a], Dewa Made Priyantha Wedagama [b], Dilum Dissanayake [a,*]

[a] School of Engineering, Newcastle University, United Kingdom
[b] Department of Civil Engineering, Udayana University, Kampus Bukit Jimbaran, Badung-Bali 80361, Indonesia

ARTICLE INFO

Keywords:
Road safety
Driver perception
Multiple correspondence analysis
Principal component analysis
Cluster analysis
Logistic regression

ABSTRACT

This study investigates drivers' attitudes towards road safety in Jakarta and Hanoi. A comprehensive analysis of two datasets was achieved by multi-faceted analytical approaches of Multiple Correspondence Analysis (MCA), Cluster analysis (Hierarchical and Two-step) (CA), Principal Component Analysis (PCA) and Multinomial Logistic Regression (MLR). MCA and CA were used to segment the drivers based on their sociodemographic characteristics and driving related information where three clusters were identified from each dataset. There were 60+ attitudinal statements in the datasets, therefore dimension reduction process was conducted using PCA. Five Principal Components were generated from each dataset, four of them common to Jakarta and Hanoi (*"Safe driving practices and behaviour", "Road safety enforcement and education", "Driver behaviour at signalised junctions", "Road infrastructure and roadside facilities"),* the fifth component being unique to each city (*"Road infrastructure design issues"* for Hanoi and *"Type of motorised vehicles"* for Jakarta). Finally, MLR was used to investigate how the level of perception on the identified components varied across sociodemographic clusters in each city. Employed young adults in Jakarta perceive that "*Driver behaviour at signalised junctions*" directly influences road safety compared to teenage and young adults who are in education. The reverse is true for Hanoi. Employed young drivers in both cities are less likely to recognise the importance of "*Road safety enforcement and education*" compared to teenage and young drivers in education. The study shows that the perception of road safety among drivers varies based on various factors including their sociodemographic traits and driving experience.

1. Introduction

According to the World Health Organisation (WHO), Road Traffic Crashes (RTCs) remain a serious public health and economic burden in the 21st century, and current trends indicate that this will continue for the foreseeable future (WHO, 2018a). It was reported that 1.35 million people on average die annually due to RTCs, with about 50 million others injured. RTCs have caused significant socio-economic costs accounting for about 3% of the global Gross Domestic Product (GDP), with the figure being even higher (5%) in high in low- and middle-income countries (Sam et al., 2019).

There was a significant relationship between the risk of road traffic deaths and income level of countries, with the risk 3 times higher in low-income countries (Gross National Income (GNI) ≤ $1,025), for example, with an average rate of 27.5 deaths per 100,000 population, than in high-income countries (GNI ≥ $12,376) where the average rate is 8.3 deaths per 100,000 population (WHO, 2018a). Moreover, low- and middle-income countries accounted for over 90% of global road traffic deaths (WHO, 2018b). According to the World Bank Country Classification (WB, 2021), South-East Asia is categorised as either low-income or middle-income economies. The fatality rate in South-East Asia was reported as very high as 20.7 deaths per 100,000 population (WHO, 2018a).

Most road traffic deaths in South-East Asia are among vulnerable road users (VRUs), with motorcyclists accounting for about 43% (WHO, 2018a). Bui et al. (2020) claimed that motorcyclists and their passengers accounted for 90% of road crashes in Vietnam, and motorcycling was responsible for 75% of economic losses ($2 billion per year). The high RTC rates in low-income countries have been attributed to growth in motorised vehicles, higher number of people killed per crash, poor

* Corresponding author.
E-mail address: dilum.dissanayake@ncl.ac.uk (D. Dissanayake).

https://doi.org/10.1016/j.ssci.2022.105869
Received 12 September 2021; Received in revised form 22 May 2022; Accepted 27 June 2022
Available online 28 July 2022

public health structure and poor enforcement of traffic laws and regulations (McIlvenny et al., 2004), coupled with lack of resources and innovation to address road safety concerns, including infrastructure, policy formulation and implementation, and capacity building (Chekijian, 2012). McIlvenny et al. (2004) further noted that RTCs were largely caused by multiple factors including road infrastructure, vehicles and human factors, with improper human behaviour accounting for about 64–90% of casualties acting as either drivers, motorcyclists, pedestrians or cyclists. Moreover, inappropriate driver behaviour in the form of violation of traffic rules and unsafe driving practices was widely considered to be a major factor in over 90% of road accidents (Li and Tay, 2013). Kayani et al. (2011) added that high risk human behaviours that were significantly correlated with RTCs included over-speeding, driving under the influence of alcohol and drugs, distracted drinking and neglect of protective gear. Outcomes of these studies confirm that safe interaction of infrastructure, vehicles, and road users will be required to achieve the goal of reducing the risk of RTCs. Whereas there has been significant progress in incorporating road safety issues in the design, construction and operation of new roads and vehicles, road safety managers and researchers across the world still have limited knowledge and coordinated understanding of human factors due to the complexity and dynamic nature of human behaviour.

Substantial amount of research has been conducted on the attitudes, beliefs and perceptions that contribute to behaviour of road users in high-income countries to provide recommendations aimed at safer road use. However, there is limited research thus far in this context attributed to low-income countries (Kayani et al., 2011; Wedagama and Dissanayake, 2020). Although considerable attention is required to improve road infrastructure and vehicle roadworthiness in those low-income countries, understanding human factors is critical in achieving road safety targets (Kayani et al., 2011). Papadimitriou et al. (2013) revealed that road safety improvement largely depends on actual road user behaviour, which is based on users' attitudes and perceptions, and therefore, knowledge of opinions and beliefs may aid in understanding road user behaviour. The overall aim of this research is therefore to conduct a study on drivers' perception of road safety in South-East Asia, with particular focus on the cities of Jakarta and Hanoi as case studies.

This will be achieved by three main objectives: 1) to investigate drivers' perception towards road safety and identify the main elements (also known as factors/components) using Principal Component Analysis, 2) to identify clusters or segments of drivers with similar socio-demographic traits such as location, age, occupation, driving experience and driving frequency, using Multiple Correspondence Analysis coupled with Hierarchical and Two-step Clustering approaches, and 3) to explore the relationship between the identified driver clusters and the components of their perception of road safety, using Discrete Choice Analysis methods. Research on road user perception and awareness of road safety will aid policy makers to understand their needs and provide potential inputs and constructive recommendations for future policies and strategic decisions.

This paper begins by providing a critical review of previous research carried out in the road safety domain focusing on RTCs and factors influencing them as well as methods of data collection. This is followed by a description of the data collection approach used for the study. Furthermore, the methodological approach for data analysis and results of the analysis are presented. Finally, the paper provides a detailed discussion of the research findings in comparison with previous research.

2. Literature review

According to the Highway Safety Manual (HSM) of the American Association of State Highway and Transportation Officials (AASHTO), road crash risk factors have been categorised into human, vehicle and road environmental descriptions, with each contributing 93%, 34% and 13% to RTCs, respectively (Ahmed, 2013). Most of the research in road safety in the past few decades focuses on these main three descriptors. Wang et al. (2002), in one of the early research projects, pointed out that driver perception of safety in the driving environment has a great influence on driving behaviour and task performance. They clearly stated that only a small number of studies have attempted to measure the perception of safety. Speck (2012) further confirms that both actual and perceived risks are essential when achieving road safety targets in which safety is not just about being safe but also feeling safe. Recently published road safety research emphasised the importance of considering both actual and perceived safety risks for policymaking (Wedagama et al., 2020; Akgün-Tanbay et al., 2022).

Inappropriate road user conduct was a major contributor to road crash risks, with about 70–80% of RTCs attributed to drivers' fault in most high-income countries (de Oña et al., 2014). Even though 95% of RTCs were attributed to road user conduct, Sayed et al. (1995) and Li and Tay (2013) revealed that more than 90% of those were caused by drivers' error. It can therefore be concluded that driver behaviour is the main human related contributor to RTCs.

The WHO revealed that VRUs accounted for more than 50% of road traffic deaths, yet they contributed less than 5% to RTC causations (WHO, 2018). There is high disability rate as a result of road traffic injury in low-income countries due to the high proportion of pedestrians, cyclists and motorcyclists, in other words VRUs, involved in RTCs (Zafar et al., 2018; Widyastuti et al., 2007). Sivasankaran and Balasubramanian (2020) further revealed that the disproportionately higher rate of VRUs involved in RTCs in Asian countries was due to excessive speeding on multilane roads, lack of physical separation, lack of formal crossing facilities and improper lighting. Recent research in Namibia, one of the most dangerous countries in Sub Saharan Africa in terms of road safety, confirms that RTCs involving VRUs occur under both intersection and non-intersection conditions and often result in serious injuries (Jones et al., 2020).

Low- and middle-income countries are undergoing intensive development in intra-urban as well as inter-urban contexts. Sobngwi-Tambekou *et al.* (2010) reported that road traffic deaths per 100 million kilometres driven on interurban roads in low-income countries is more than 35 times higher than on similar roads in Europe or United States of America. This has been attributed to unsafe driver behaviour through excessive speeding and dangerous overtaking manoeuvres as well as speed differences among cars, trucks, buses and motorcyclists. Additionally, Fruhen and Flin (2015) revealed that hostility and lack of consideration among different road users, for example car drivers towards motorcyclists and cyclists, would be a major cause of RTCs on multilane roads.

Although vehicle factors contributed a comparatively smaller percentage to RTCs compared to human factors as per AASHTO guidelines, the high rates of fatalities and injuries from RTCs in low-income countries has been attributed to the rising volume of motorisation, dominated by motorcycles, and dependence on importation of older and "second-hand" vehicles from high-income countries (Hazen and Ehiri, 2006; McIlvenny et al. (2004). This is exacerbated by lack of regulation, poor inspection on border entry points, lack of maintenance and low-quality spare/substitute parts, especially tyres and brakes (Chen, 2010).

Previous research found that road environmental factors, such as weather conditions (snowfall, fog, wind and temperature) influence the occurrence of RTCs (de Oña et al., 2014; Schlögl, 2020; Theofilatos and Yannis, 2014), with precipitation being the most analysed meteorological variable in terms of impact on road safety. Akgün et al. (2018), Atombo et al. (2016) and Edquist et al. (2009) also noted that drivers' behaviour was dependent on the road environment, for example road geometry, road signs and markings, lighting conditions and weather conditions. They claimed that road surface conditions were the most influential factors that affect driver perception of speeding and lane changing. There were other underlying socio-economic factors such as high poverty levels as well as inadequate surveillance systems, public health infrastructure, emergency medical services and trauma care

Table 1
Comparison of highly motorised countries in South-East Asia.

Country	% of households that have a motorcycle [a]	% of households that have a car [a]	Income level [b]
Indonesia	85	4	Lower middle
Malaysia	83	82	Upper middle
Vietnam	87	2	Lower middle
Thailand	86	51	Upper middle

Sources: Lee (2016)[a], World Bank (2021)[b].

Table 2
Themes and subthemes in the questionnaire survey.

€Theme	Sub-themes
Driving Safety Tips	What practices do you usually observe for traffic safety? What do you usually do to avoid traffic accidents?
Driving through Intersections	When the traffic signal turned into YELLOW from RED, what are you supposed to do? Have you ever ignored traffic signals? What is the likely reason of your ignoring of the traffic signals?
Traffic Accidents	In your opinion, where is one most likely to have a traffic accident while driving? In your opinion, who are the main causes of traffic accidents while driving? In your opinion, what is the major cause of traffic accidents while driving? In your opinion, at what time are traffic accidents most likely to happen while driving?
Traffic Safety	In your opinion, what measures can be taken to reduce or avoid Traffic accidents while driving? In your opinion, what is an important environment for safety traffic in your city? Which group needs to have traffic education?

systems that contribute to RTCs, especially in low-income countries, as informed by research (Hazen and Ehiri, 2006).

Previous research investigated the connection between RTC risks and other influential driver related factors including driving experience (i.e. licence status, years of driving, past accident involvements and distance driven), driver's sociodemographic characteristics (i.e. gender, age, income, commuter status, level of education and marital status), drivers physiological and psychological state, personal traits, social cognition (attitudes and risk perception), individual emotional state, cultural settings and values (fatalism and superstition), use of alcohol and drugs, use of seatbelts, violation of speed limits, dangerous manoeuvres, ignoring of posted signs and use of cell phones (Akgün et al., 2021; Arnau-Sabatés et al., 2012; Atombo et al., 2016; de Oña et al., 2014; Dissanayake, 2004; Kayani et al., 2011; Şimşekoğlu et al., 2013; Ulleberg and Rundmo, 2003; Vanlaar and Yannis, 2006; Wang et al., 2002; Wedagama and Dissanayake, 2009; Wedagama et al., 2011; Wedagama et al., 2008; Yannis et al., 2005). Furthermore, Yannis et al. (2005) indicated that parameters usually considered in the general framework of travel choices, for example trip duration and trip cost, also had a significant impact on drivers' choices towards accident risk reduction alternatives. Human incapacitation and driver fatigue were also highlighted by Chen (2010) as major accident contributing factors.

Data collection for most attitudinal and behavioural research in road safety has been conducted based on questionnaire surveys, face-to-face interviews, telephone surveys, direct observations, or through more integrated methods (Papadimitriou et al., 2013). Traditionally, drivers' attitudes and perception of accident risks have always been investigated through surveys with data gathered through "self-reported risky driving behaviour" (de Oña et al., 2014). For example, Deffenbacher et al. (1994) high-income a Driver Anger Scale to measure anger induced by traffic situations resulting from hostile gestures, illegal driving, police presence, slow driving, discourtesy and traffic obstructions; Lajunen and Summala (1995) used a Driving Skill Inventory to assess drivers' skill and safety, while Reason et al. (1990) used questionnaires to measure driver behaviour in forms of errors, lapses and violations. Questionnaires, both self-administered or interviewer administered, are preferred for data collection since they are cheaper, less time consuming, resilient to bias and provide completed responses.

In conclusion, it has been well documented by previous research that RTCs are multidimensional and random occurrences that involve interaction among infrastructure, vehicle and human related factors. Drivers' perceptions have been considered an important influence on their driving behaviour and road safety (Wang et al., 2002). However, research on RTCs in low-income countries, particularly in South-East Asia, is very limited due to lack of reliable data, limited funding and inadequate technical capacity, among other factors. Therefore, this study will focus on driver perception towards road safety using the data collected via a face-to-face road safety awareness questionnaire survey in Jakarta and Hanoi.

3. Data and methodology

3.1. Selection of case study areas

The high proportion of RTCs in low- and middle-income (LMICs) countries has been partly attributed to growth in motorization dominated by motorcycles, especially in urban centres (Senbil et al., 2007). Motorcycle ownership in Indonesia, Malaysia, Thailand and Vietnam has risen to high levels compared with the rest of the world (Senbil et al., 2007; Kitamura et al., 2018; Lee, 2016). Table 1 below summarises vehicle ownership and income in the countries mentioned.

This study focuses on Indonesia and Vietnam as they share similar levels of motorisation (Lee, 2016) as well as they belong to lower middle-income category as defined by the World Bank (WB, 2021). The capital cities in Indonesia (Greater Jakarta Region) and Vietnam (Hanoi) were used as case study cities. Jakarta is the most populated city in South-East Asia. There were about 16.2 million motorcycles and over 3.4 million private cars in Jakarta in 2020 (Indonesia Police Office, 2022). Like Jakarta, Hanoi is also facing high rates of motorisation, which has worsened traffic congestion and RTCs. The ratio of motorcycles to cars on Hanoi's streets is 10:1 and 70% of RTCs in the city were associated with motorcycle use (Turner, 2020). Therefore, Jakarta and Hanoi provide a fair representation of road safety issues in South-East Asia with particular attention to the lower middle-income contexts. The surveys were conducted with attention to urban areas that represent highest burden of RTCs due to high traffic volumes and density.

3.2. Data collection

The data from Jakarta and Hanoi was collected via questionnaire surveys focusing on motorised vehicle drivers. The questionnaire was composed of two parts. The first part consists of 10 questions including the district that the respondent resides, and their sociodemographic characteristics such as household information (income, vehicle ownership), individual's information (gender, age, occupation) and driving related information (experience, driving frequency, type of vehicle they drive and whether she/he had any traffic safety education). The second part was designed to examine drivers' awareness of road safety with four themes: "safe driving tips", "driving through intersections", "road traffic accidents" and "road traffic safety" (Table 2). Each theme had 2–4 subthemes where 5–10 statements were included in them to gather respondent's agreement or disagreement (via YES/NO responses).

The questionnaire was translated into respective local languages for ease of use followed by back-translation to find any mistakes before finalising them. Four graduate students from each case study city were recruited and trained for the data collection activity. The training

Table 3
Population and sample statistics.

Greater Jakarta Region		Population [a]		Sample	
		#	%	#	%
Central Districts	North	1,716,591	10%	31	8%
	East	2,916,020	17%	67	17%
	South	2,246,140	13%	52	13%
	West	1,276,097	8%	37	9%
	Central	924,690	5%	31	8%
Other Districts	Bogor	1,081,009	6%	34	8%
	Depok	1,809,120	11%	24	6%
	Tangerang	2,139,891	13%	60	15%
	Bekasi	2,873,484	17%	69	17%
	Total	*16,983,042*		*405*	
Hanoi Region		Population [b]		Sample	
		#	%	#	%
Central Districts	Ba Dinh	221,893	6%	31	8%
	Hoan Kiem	135,618	4%	28	7%
	Tay Ho	160,495	4%	24	6%
	Thanh Xuan	293,524	8%	39	10%
	Cau Giay	292,536	8%	39	10%
	Dong Da	371,606	10%	35	9%
	Hai Ba Trung	303,586	8%	24	6%
Other Districts	Hoang Mai	506,347	14%	47	12%
	Long Bien	322,549	9%	31	8%
	Nam Tu Liem	264,246	7%	28	7%
	Bac Tu Liem	335,110	9%	39	10%
	Ha dong	397,854	11%	28	7%
	Total	*3,605,364*		*393*	

[a] Population as of 2017–18 (Jakarta Statistics, 2022; Depok Statistics, 2022; Bogor Statistics, 2022; Bekasi Statistics, 2022; Tangerang Statistics, 2022).
[b] Population as of 2019 (General Statistics Office in Vietnam, 2019).

session included several themes such as introducing the aim of the survey, debriefing, and question-by-question explanation. In some occasions, the respondents will require some additional clarifications on questions while completing the survey. Therefore, the survey staff training activity was crucial. As part of the university practice, the ethical approval of the questionnaire was secured before piloting it in Jakarta and Hanoi.

The pilot surveys were conducted in both cities in May 2017 where 40 samples were collected from Jakarta and Hanoi (10% of the targeted sample). The pilot survey was a useful exercise to identify any issues with the questionnaire design, questions and difficulties related to understanding to minimise any biases. During the pilot survey, the survey staff in each city was supervised by two post-graduate students from the UK who speak Indonesian and Vietnamese languages. Taking on board the feedback received from the piloting exercise, the questionnaires were improved to be used for the data collection. The main surveys were conducted in June-July 2017 where the respondents filled the questionnaires in face-to-face communication with interviewers.

According to Ortúzar and Willumsen (2011), it was identified that the minimum sample size required was 384 at 95% confidence level from each case study region. Accordingly, 405 and 393 questionnaires were completed in Jakarta and Hanoi, respectively. As driver population statistics, for example age, gender, and driving experience were not available in Jakarta and Hanoi, stratified random sampling technique was applied taking into consideration of population statistics in district levels (Table 3). This was to ensure that the sample reflects the diversity of the population across the region. Each district (strata) was then considered individually when applying random sampling. The interviewers organised to interview respondents at their residences and business premises in urban areas especially to target car/van/minibus drivers. In contrast, interviews were conducted at respondents' residences in suburban areas aiming for motorcycle riders. This was based on previous research that high income households tend to reside in urban areas and low income households choose to live in suburban areas in Jakarta and Hanoi (Janssen et al., 2021; Gubry et al., 2009). Interviewing respondents at their residences and business premises was considered more appropriate compared to roadside interviews due to road safety issues, e.g. congested traffic and chaotic roadside environment, and health reasons, e.g. poor air quality due to heavy traffic.

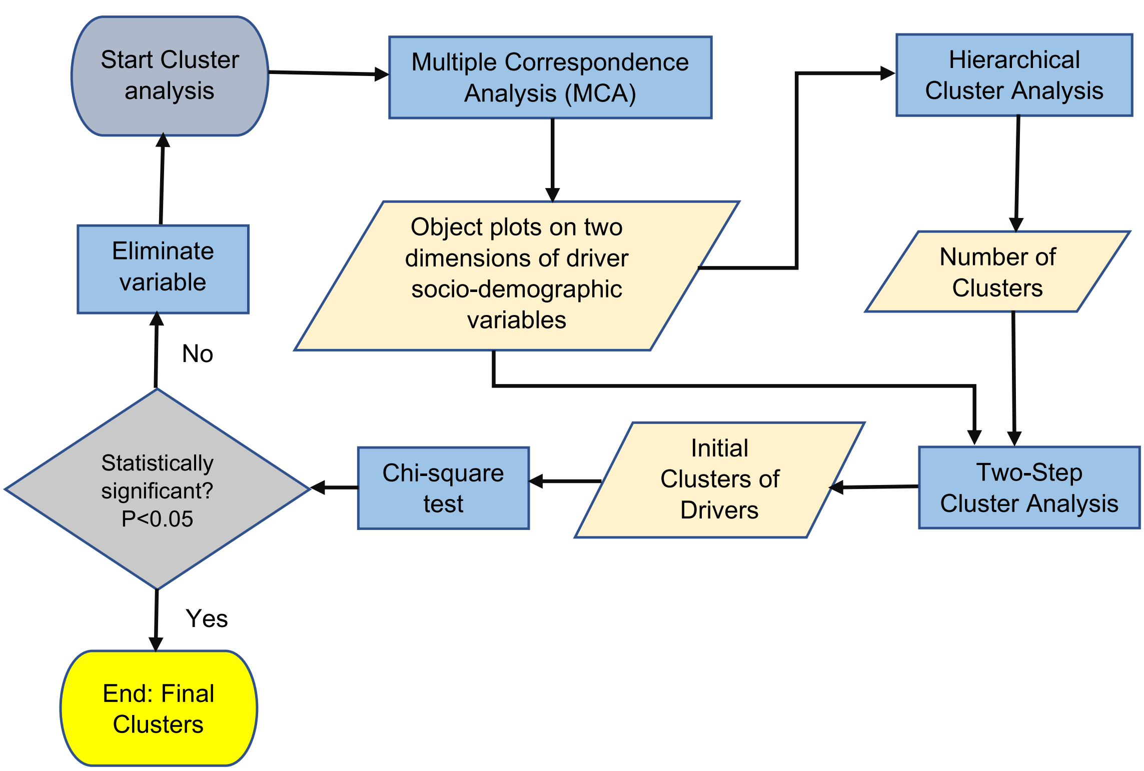


Fig. 1. Flowchart of the data clustering process.

Table 4
Descriptive Statistics of Sociodemographic Variables.

	Category	Jakarta, N = 405		Hanoi, N = 393	
		Frequency	Proportion (%)	Frequency	Proportion (%)
Gender	Male	258	**64**	218	**55**
	Female	147	36	175	45
Age	≤19	67	17	53	13
	20–29	261	**64**	236	**60**
	30–39	44	11	61	16
	40–49	21	05	33	08
	50 ≥	12	03	10	03
Occupation	Public Employee	27	07	46	12
	Private Employment	192	**47**	111	**28**
	Self Employed	33	08	30	08
	Housewife/husband	05	01	10	03
	Unemployed	01	00	5	01
	Retired	9	02	16	04
	Student	138	**34**	172	**44**
Driving Experience	<1 year	45	11	25	06
	1–5 year(s)	109	**27**	153	**39**
	5–10 years	159	**39**	96	**24**
	10–20 years	63	16	34	09
	>20 years	29	07	08	02
Driving Frequency	Everyday	270	**67**	263	**67**
	2 or 3 times a week	75	19	42	11
	once a week	14	03	03	01
	once a month	02	00	00	00
	hardly ever	44	11	08	02
Type of Vehicle	Motorcycle	237	**59**	258	**66**
	Sedan/SUV	152	38	58	15
	Bus/Minibus	14	03	77	20
	Others	02	00	00	00
Safety education	Yes	86	21	297	**76**
	No	319	**79**	96	24

3.3. Methods used for data analysis

The main data analysis was conducted using SPSS software platform and it involved grouping drivers with similar demographic characteristics together (by data clustering) and then determining the components generating from the attitudinal statements (dimension reduction by Principal Component Analysis) and level of each cluster's perception towards road safety (by Logistic Regression analysis). The selected research approaches are described in detail with attention to their suitability for data analysis in this study.

3.3.1. Clustering of sociodemographic data

The respondents were clustered into similar and easily manageable groups based on their sociodemographic characteristics (individual and household information) as well as the variables of their driving related characteristics (experience, driving frequency, type of vehicle they drive and whether she/he had any traffic safety education) collected via the questionnaire surveys. This was done using Multiple Correspondence Analysis (MCA) in combination with hierarchical and two-step clustering approaches in an iterative process (Fig. 1).

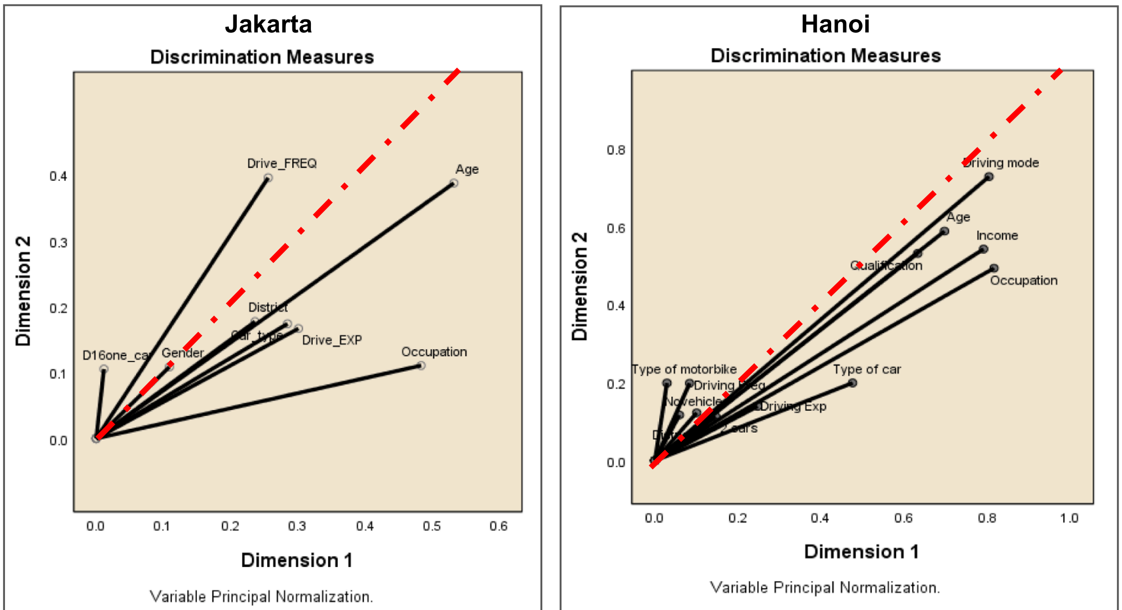
MCA is a powerful tool widely used in social science research in recent years. It explores all types of variables in datasets (nominal, ordinal, continuous), indicating if any relationship exists between variables and how they are related, and offering statistical results that can be seen both analytically and visually (Ali et al., 2018; Costa et al., 2013). MCA was chosen in this research to visualise the patterns geometrically by locating each variable/unit of analysis as a point in a low-dimensional space, in other words two-dimension graphs. Visual illustrations provide a simplified way to understand and interpret large and complex datasets by effectively summarising and structuring the relations among variables (Das and Sun, 2016; Ali et al., 2018). The number of dimensions to be used is usually determined as appropriate to the analysis undertaken, but two-dimensional comparison is deemed sufficient for easy visualisation (Greenacre and Blasius, 2006).

Once the MCA analysis was completed, hierarchical clustering was conducted to determine the number of clusters using the continuous object points from MCA as inputs. Dendrograms were used to determine the number of clusters in the data sets. When analysing a dataset with both continuous and ordered variables, previous research have used two-step clustering straight away instead of using MCA and hierarchical clustering. However, two-step clustering process determines the number of clusters automatically without giving an opportunity to investigate the most appropriate number of clusters based on the relationships of similarity among groups. Carrying out hierarchical clustering will help select the number of clusters by observing the dendrograms. When conducting hierarchical clustering in this research, Ward's method and squared Euclidean distance were adopted for measuring the similarity and distance between clusters, respectively. Ward's method was considered suitable because it offered more compact and distinct clusters (Hair et al., 1998). The squared Euclidean distance was preferred because it offered easy computational analysis compared to the basic Euclidean distance measure.

The two-step cluster analysis was carried out to classify the clusters using object scores from MCA as suggested by previous research (Ribbens et al., 2008; Yoshida et al., 2020) and the number of clusters determined by hierarchical clustering. The quality of the clusters was evaluated by their silhouette measures, the cluster size ratios (largest cluster to smallest cluster) as well as chi-square tests.

3.3.2. Principal Component Analysis

Principal Component Analysis (PCA) was performed to reduce the large set of driver safety perception variables, in other words attitudinal statements, to a smaller set of dimensions that could be easily and meaningfully analysed. The first step in PCA was to conduct initial checks, which included sample size adequacy and correlations between variables. Field (2017) revealed that the reliability of PCA depends on sample size, component loadings and communalities. It is advisable that for a sample of 300 or more to provide a more stable solution, there is a need to have enough variables to adequately measure all the components. The sample size adequacy was checked using the Kaiser-Meyer-



Dimension	1	2	Mean
Total Eigen Value	2.21	1.62	1.92
% of Variance	27.66	20.32	23.98
Cronbach's Alpha	0.63	0.44	0.55

Dimension	1	2	Mean
Total Eigen Value	4.88	3.97	4.43
% of Variance	40.68	33.06	36.87
Cronbach's Alpha	0.87	0.82	0.84

Fig. 2. Final Discrimination Measures of the Variables.

Olkin (KMO) measure. KMO values less than 0.5 are considered inadequate while values closer to 1 are recommended (Field, 2017).

Component extraction process was then carried out to determine the number of components to be retained. Scree plots were used for this in combination with other methods including elimination of trivial factors, prior criteria and percentage of cumulative variance. Additionally, oblique rotation using promax was adopted for component rotation since it was likely that there is a correlation between components that explain drivers' perception towards road safety.

Field (2017) recommended that component loading values between 0.3 and 0.4 were acceptable, even though it was preferred to have values greater than 0.5. For this research, the lower limit of component loading was set as 0.4. Finally, the reliability test for the extracted components was conducted by calculating the Cronbach's alpha (α) value. It is claimed that α values of 0.7 to 0.8 are desirable (Field, 2017).

3.3.3. Multinomial Logistic Regression Analysis

Multinomial logistic regression models were developed to examine the association between the driver clusters and their perception of road safety focusing on the components derived from PCA. Although both linear and logistic regression techniques possess similar strengths in performing hypothesis testing to examine significant levels of model coefficients (Wright, 1995), the latter was the preferred option for this study because of its ability to analyse categorical and non-linear data. The maximum likelihood method was used to estimate parameters by selecting coefficients, which make observed values most likely to have occurred (Field, 2017). If the dependent variable has more than two categories, Multinomial Logistic Regression (MLR) should be employed. The cluster results from the two-step cluster analysis were used as dependent variables, whereas the components from PCA were taken as covariates, which are the predictors (independent variables) of the drivers' perception of road safety. It was noted that while conducting MLR analysis, one category must be regarded as a reference for comparison purposes (Field, 2017).

The goodness of fit of the model was assessed by comparing observed and predicted values of the outcome using the log-likelihood statistic, deviance statistic and pseudo R^2 (Field, 2017).

4. Results

4.1. Sample description

Descriptive statistics related to the sociodemographic variables of the collected data sample are summarised in Table 4. Accordingly, both cities were dominated by male young adult drivers whose daily commute is for either private employment or education. Male drivers as well as young novice drivers are more susceptible to road crash involvement due to over speeding, distracted driving and driving under the influence of alcohol (Borowsky and Oron-Gilad, 2013; Cordellieri et al., 2016). Additionally, most drivers in both data samples drive daily and have driving experience of 1–10 years. The frequency and severity of road traffic crashes depends on driving experience and frequency. Furthermore, Bui et al. (2020) revealed that motorcycles accounted for the biggest percentage of motorised vehicles in South-East Asia and were largely responsible for the high road crashes. It is shown that most drivers in Jakarta and Hanoi use motorcycles as a means of transport.

Sami et al. (2013) and Borrell et al. (2022) indicated that the level of education of road users influenced occurrence of RTCs. These studies noted that lack of road safety education in low-income countries could be leading to higher RTC rates. Table 4 indicates that most of the respondents in Hanoi had received road safety education in their schools compared to Jakarta. This implies that road safety programmes are incorporated in the curriculum of schools in Vietnam.

4.2. Identification of clusters of driver population

4.2.1. Multiple Correspondence Analysis

Multiple Correspondence Analysis (MCA) was conducted to graphically identify the association among variables and transform categorical data into continuous data for hierarchical cluster analysis. It is reported

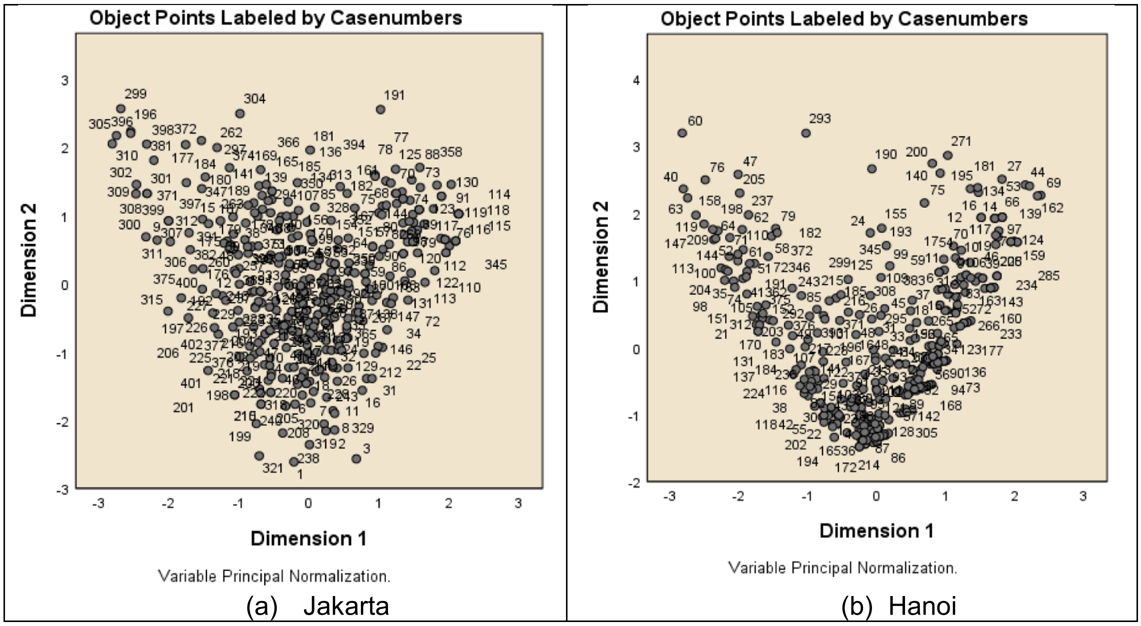


Fig. 3. Final MCA Object Points in Two Dimensions.

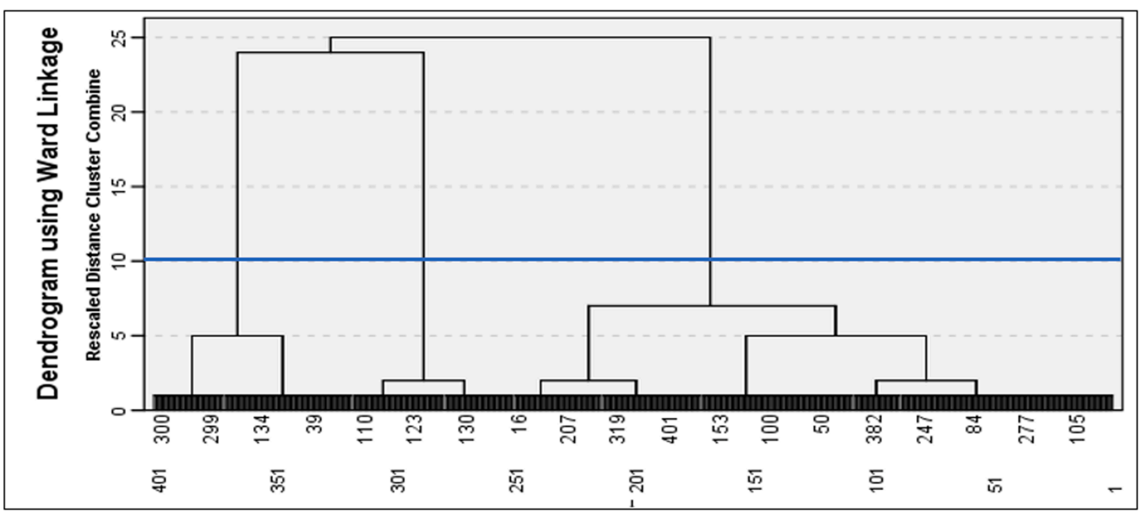


Fig. 4a. Dendrogram as a result of Hierarchical Clustering – Jakarta.

that dimensions for Hanoi provided a high measure of reliability (α = 0.84) whereas the Jakarta sample had moderate reliability (α = 0.55) (Fig. 2). The moderate reliability was due to reduced number of questions attached to dimension 2 as well as differences in scales (gender has two levels compared to driving frequency that has 5 levels). α values smaller than 0.7 was considered acceptable in exploratory research (Johnson and Wichern, 2007). The dimensions for Hanoi explained relatively high percentage of variance compared to Jakarta (73.4% Vs 48.5%). Hair et al. (2014) claimed that 95% of variance is most appropriate for natural science research, 60% or less was considered satisfactory in social sciences. Therefore, both MCA models were considered as reliable. MCA enabled discrimination measures of every variable onto each dimension to be visually presented (Fig. 2).

The graphical plots provided a visual representation of the correlation between variables and discrimination of each variable onto each dimension whereby a red 45°-line drawn from the origin was used to the show degree of correlation (Costa et al., 2013).

It indicates that for Jakarta, dimension 1 is more correlated with location, gender, age, occupation, driving experience and vehicle type, whereas dimension 2 is associated with gender, household with one car and driving frequency. In Hanoi, dimension 1 is represented by age, occupation, income, educational qualification, driving mode, type of car and driving experience, while dimension 2 is associated with location, household with no vehicle, type of motorbike and driving frequency.

4.2.2. *Cluster analysis*

Hierarchical cluster analysis was performed to determine the number of clusters. The continuous object points from MCA were used as inputs for this analysis (Fig. 3).

The dendrograms from hierarchical cluster analysis are appended in Fig. 4a and Fig. 4b where three clusters of drivers were extracted from each case study.

Two-step cluster analysis was then undertaken to group drivers with similar sociodemographic characteristics together with object scores from MCA and the number of clusters determined by hierarchical clustering. Good quality clusters of drivers were produced since their silhouette measures of cohesion and separation were above 0.5 which is recommended by Norusis (2011) as desirable. It was further shown that

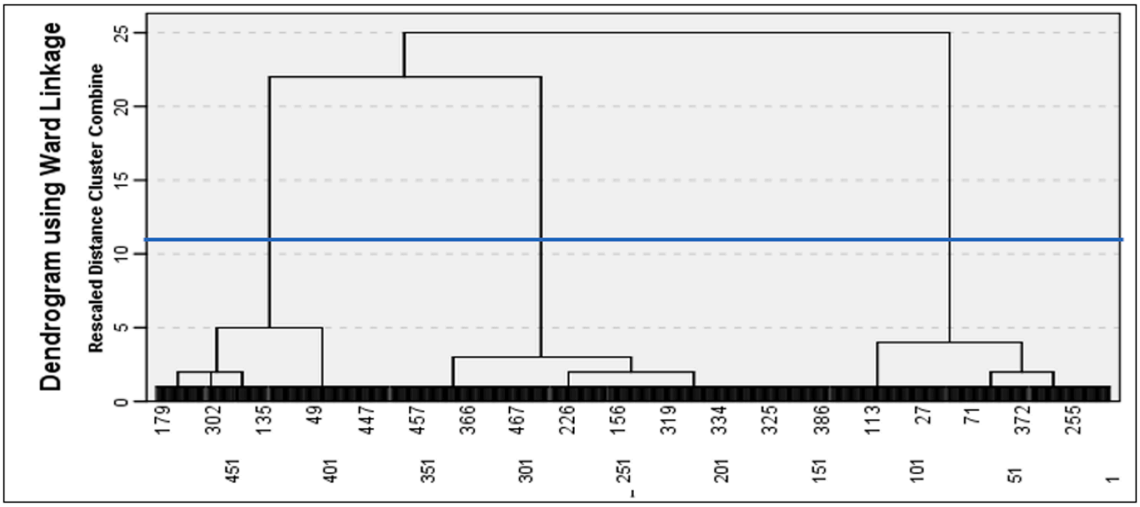


Fig. 4b. Dendrogram as a result of Hierarchical Clustering – Hanoi.

Table 5
Summary of Results of the Cluster Analysis.

	Jakarta		Hanoi	
Cluster Analysis				
# clusters	3		3	
Cluster quality	Good		Good	
Ratio of largest to smallest cluster	2.92		2.75	
Cluster Details	***Cluster Name***	***Proportion***	***Cluster Name***	***Proportion***
Cluster 1	*Young, employed adults driving car or motorcycle*	57.5%	*Young and middle aged employed with postgraduate qualifications driving car*	20.4%
Cluster 2	*Young and middle aged, employed males driving car or motorcycle*	22.9%	*Teenage and young adults in education riding motorcycles*	23.5%
Cluster 3	*Teenagers in education riding motorcycles*	19.7%	*Young adults either employed or in education with undergraduate qualification riding motorcycles*	56.1%

the data samples exhibited an acceptable size ratio of largest cluster to smallest cluster (2.92 and 2.75 for Jakarta and Hanoi respectively). Furthermore, chi-square tests were conducted, which revealed that all the variables within each cluster solution were statistically significant at 95% confidence level. The multiple comparison problem was checked using the Bonferroni correction and the level of significance was tested at $P < 0.005$ since 10 demographic variables were considered in our analysis (Bradburn, 2022). Those variables include residential Location, Age, Gender, Occupation, Monthly household income, Household vehicle ownership, Type of vehicle, Driving frequency, Driving experience, and Traffic safety education.

Table 5 shows that three clusters were identified from each dataset of Jakarta and Hanoi. Respondents in Jakarta were clustered into "Young employed adults driving car or motorcycle", "Young and middle age employed males driving car or motorcycle" and "Young adults in education riding motorcycles"; whereas drivers in Hanoi were grouped into "Young and middle age employed with postgraduate qualifications driving car", "Teenage and young adult in education driving motorcycles" and "Young adult either employed or in education with undergraduate qualification riding a motorcycle". The details of the driver cluster classifications are provided in Appendix A**.**

The relatively similar clusters identified in the two cities suggest that road users in South-East Asia have similar sociodemographic traits and may exhibit related behaviour. According to previous research, South-East Asia is dominated by young adult drivers who ride mainly motorcycles. This resonates with Soehodho (2009) who stated that the most vulnerable age group in Indonesia in terms of road crashes was aged between 16 and 30 years. The study claimed that young drivers had high chances of being involved in road crashes because they lacked driving experience and were ignorant of the impacts of risky driving behaviour. The composition of the clusters in both cities coincides with the motorisation statistics of Vietnam and Indonesia, which revealed that motorcycles accounted for about 94% and 81% of motorised vehicles, respectively. The high rates of road accidents among motorcyclists were attributed to the large number and proportion of motorcycles, inherent dangerous characteristics of the traffic environment and aberrant riding behaviour such as over speeding and careless lane shifting/overtaking (Bui et al., 2020; Tuan, 2015).

The study also revealed that there was a significant proportion of student drivers in South-East Asia who commute daily to school/university using motorcycles. This cluster belongs to the road user group (5–29 years) that is the most killed in the world as a result of road traffic injury (WHO, 2018b). Borowsky and Oron-Gilad (2013) revealed that young novice drivers were more prone to crash involvement than experienced drivers due to distracted driving and exceeding speed limits. Furthermore, the study also identified a cluster of male middle-aged drivers in Jakarta and Hanoi whose daily commute was by either car or motorcycle. This male dominated cluster represents the gender that is most prone to risky driving behaviour. According to Cordellieri et al. (2016) and Oltedal and Rundmo (2006), the rate of males' involvement in fatal accidents was twice as high as that of females because men were more susceptible to over-speeding, traffic law violation and driving under the influence of alcohol and drugs.

Table 6
Output of the Component Structure from PCA for Jakarta and Hanoi.

Statement	C	Com	α
(a) PCA results – Jakarta			
Component 1: Safe Driving Practices and Behaviour			0.82
Reducing speed before a sharp bend	*0.81*	*0.65*	
Drive slowly or stop before approaching intersections	*0.77*	*0.63*	
Turning on the headlights before sunset	*0.74*	*0.51*	
Drive while fastening your seat belt	*0.69*	*0.48*	
Driving/riding within your lane and not overtaking	*0.64*	*0.48*	
Stop to allow pedestrians to cross at zebra crossing	*0.64*	*0.50*	
Not driving after drinking alcohol	*0.57*	*0.33*	
Component 2: Road Safety Education and Enforcement			0.84
Do not drive/ride after drinking as a measure	*0.67*	*0.63*	
Speed limit enforcement as an environment for safety	*0.65*	*0.37*	
Driving in within the specified speed limit as a measure	*0.63*	*0.33*	
Fastening seatbelts at all times while driving	*0.61*	*0.56*	
On-street parking enforcement as an important environment	*0.57*	*0.47*	
Stop on-street parking as a measure	*0.54*	*0.53*	
Do not use cell phone while driving/riding as measure	*0.53*	*0.52*	
Dissemination of traffic education	*0.52*	*0.29*	
Maintaining one's lane/not overtaking	*0.50*	*0.45*	
Component 3: Road Infrastructure and Roadside Facilities			0.70
Narrow streets/roads as major accident cause	*0.68*	*0.43*	
On-street parking as major accident cause	*0.68*	*0.49*	
Lack of streetlights as major accident cause	*0.60*	*0.43*	
Bad road surface as a major accident cause	*0.51*	*0.44*	
minor or narrow roads as an accident hotspot	*0.50*	*0.26*	
Running out into the road suddenly (pedestrian)	*0.50*	*0.44*	
Component 4: Type of Vehicles			0.59
Car as main causer of accidents	*0.68*	*0.46*	
Large-sized vehicle (truck, trailer) as main causer of accidents	*0.61*	*0.46*	
Bus/minibus (Taxi) as main causer of accidents	*0.57*	*0.41*	
Motorbike as main causer of accidents	*0.54*	*0.34*	
Major or wide roads as accident hotspots	*0.52*	*0.25*	
Component 5: Driver Behaviour at Signalised Junctions			0.59
Ignoring traffic signals	*0.84*	*0.72*	
No traffic police was present	*0.63*	*0.42*	
Was in hurry as reason for ignoring traffic signals	*0.62*	*0.42*	
No other cars/pedestrian in intersections	*0.57*	*0.34*	
KMO: 0.87			
C = Component loadings; Com = communalities; α = Cronbach's Alpha			

(b) PCA results - Hanoi Statement	C	Com	α
Component 1: Safe Driving Practices and Behaviour			0.26
Do not drink alcohol while driving as safe behaviour	*0.71*	*0.52*	
Minor or narrow roads as accident hotspot	*0.45*	*0.51*	
Careless or inattentive driving/riding as reasons for collisions	*0.43*	*0.26*	
Component 2: Road Safety Education and Enforcement			0.36
Do not drive after drinking alcohol as a safe behaviour	*0.67*	*0.46*	
No traffic police was present	*0.62*	*0.39*	
Cycle as main cause of collisions	*0.64*	*0.48*	
Don't reduce speed before approaching intersections	*0.43*	*0.41*	
Component 3: Road Infrastructure and Roadside Facilities			0.41
Bad road surface as main cause of collisions	*0.72*	*0.52*	
Improve quality of road surface as measure	*0.66*	*0.48*	
Stopping to allow pedestrians to cross at a zebra crossing	*0.51*	*0.29*	
Component 4: Road Infrastructure Design Issues			0.30
Widen the road as measure to reduce accidents	*0.68*	*0.46*	
Narrow streets/roads as a reason for collisions	*0.57*	*0.60*	
Driving/riding exceeding the maximum limit speed	*0.51*	*0.39*	
Don't transport goods exceed that limit in height and width	*0.44*	*0.34*	
Component 5: Driver Behaviour at Signalised Junctions			0.67
Traffic collision frequency	*0.88*	*0.76*	
Ignoring traffic signals	*0.86*	*0.72*	
Was in a hurry as reason for ignoring traffic signals	*0.68*	*0.43*	
KMO: 0.61			
C =Component loadings; Com = communalities; α = Cronbach's Alpha			

4.3. Components of drivers' perception of road safety

PCA was applied to attitudinal variables to identify and summarise the components of drivers' perception towards road safety. PCA is essential in identification of variable interrelatedness and elimination of redundancies (Field, 2017; Sam et al., 2019). Preliminary analysis explored the component structure and dimensionality as well as the sample size adequacy. 61 and 60 variables for Jakarta and Hanoi respectively were subjected to preliminary analysis whereby variables that contributed to components whose loadings were less than 0.4 were eliminated. The final PCA results revealed that most coefficients in the correlation matrices were above 0.3 and the KMO values were greater

Table 7
MLR Parameter Estimates.

(a) MLR results for Jakarta data

Cluster Classification (Jakarta)	Components generated from the PCA	B		Sig.
Cluster 1: *Young, employed adults driving car or motorcycle*	1. Safe Driving Practices and Behaviour	−0.14		0.38
	2. Road Safety Enforcement and Education	**−0.32**	**	0.05
	3. Road Infrastructure and Roadside Facilities	0.04		0.80
	4. Type of Motorised Vehicles	**0.47**	*	0.00
	5. Driver Behaviour at Signalised Junctions	**0.44**	*	0.00
Cluster 2: *Young and middle age employed males driving car or motorcycle*	1. Safe Driving Practices and Behaviour	−0.16		0.39
	2. Road Safety Enforcement and Education	−0.29		0.12
	3. Road Infrastructure and Roadside Facilities	0.10		0.56
	4. Type of Motorised Vehicles	**0.43**	*	0.02
	5. Driver Behaviour at Signalised Junctions	0.27		0.11

Pseudo R^2: 0.07; No. of samples: 405
Notes: Reference category: Cluster 3 (Teenagers in education riding motorcycles)
Significant factors at 95%: Bold *; Significant factors at 90%: Bold **

(b) MLR results for Hanoi data

Cluster Classification (Hanoi)	Components generated from the PCA	B		Sig.
Cluster 1: *Young and middle age employed with postgraduate qualifications driving car*	1. Safe Driving Practices and Behaviour	−0.09		0.57
	2. Road Safety Enforcement and Education	**−0.27**	**	0.09
	3. Road Infrastructure and Roadside Facilities	−0.03		0.84
	4. Road Infrastructure Design Issues	**−0.30**	**	0.05
	5. Driver Behaviour at Signalised Junctions	**−0.32**	*	0.04
Cluster 3: *Young adult either employed or in education with undergraduate qualification riding a motorcycle*	1. Safe Driving Practices and Behaviour	−0.09		0.48
	2. Road Safety Enforcement and Education	**−0.42**	*	0.00
	3. Road Infrastructure and Roadside Facilities	0.03		0.83
	4. Road Infrastructure Design Issues	−0.21		0.11
	5. Driver Behaviour at Signalised Junctions	**0.41**	*	0.00

Pseudo R^2: 0.16; No. of samples: 393
Notes: Reference category: Cluster 2 (Teenage and young adults in education riding motorcycles)
Significant factors at 95%: Bold *; Significant factors at 90%: Bold **

Table 8
Correlation between Components.

Comp.	1	2	3	4	5
(a) Component correlation matrix for Jakarta					
1	1.000	0.485	0.350	0.349	0.075
2		1.000	0.398	0.368	0.066
3			1.000	0.352	0.138
4				1.000	0.069
5					1.000
(b) Component correlation matrix for Hanoi					
Comp.	1	2	3	4	5
1	1.000	−0.197	0.135	−0.130	−0.186
2		1.000	0.074	0.201	0.233
3			1.000	−0.011	0.053
4				1.000	0.046
5					1.000

than the acceptable limit of 0.5.

Scree plots were used to extract components since the sample size was more than 300 and average communality for both data sets were less than 0.6, which is the lower limit for using Kaiser's criterion. The Scree plots were examined to determine the point of flattening out. It was observed that there was no clear break on the Scree plot. However, based on prior criterion of the structure of the questionnaire (Brown, 2009), the number of components was fixed as five. The five extracted components accounted for total variance explained of about 45% and 46% for Jakarta and Hanoi, respectively.

Oblique rotation (promax) was performed to improve the interpretability of the five extracted components. All component loadings less than 0.4 were suppressed in order to make the interpretation easier. However, the component correlation matrix showed that the extracted components for the Hanoi data were not correlated. Therefore, the process was repeated for the Hanoi data using orthogonal rotation (varimax), which is recommended for uncorrelated components (Brown, 2009). Table 6 shows that the Jakarta sample provided more reliable component loadings compared to Hanoi.

All component loadings for Jakarta were above 0.5 and each component contained more variables. Additionally, components 1 and 2 for Jakarta had high reliabilities (with α > 0.8), whereas components 3, 4 and 5 had moderate reliabilities with α values between 0.5 and 0.7. However, the sample from Hanoi provided low reliable components and number of component categories with only component 5 having a moderate reliability. The low reliability can be attributed to low correlations among the individual variables and small number of variables contributing to components.

Both datasets generated FIVE components each. Out of those, four components were common to both datasets: "**safe driving practices and behaviour**", "**road safety enforcement and education**", "**driver behaviour at signalised junctions**", and "**road infrastructure and roadside facilities**". The component "**Type of motorised vehicles**" was unique to Jakarta whereas "**road infrastructure design issues**" was related to Hanoi.

The extracted components show that drivers associated with road safety around the three dimensions of infrastructure, vehicles and human behaviour, all of which influence the occurrence and severity of road traffic accidents as well as the **3Es (Engineering, Education and Enforcement)** approach to road safety. Sami et al. (2013) revealed that road **safety education** was very significant in reducing road traffic accidents because it is through education that road users are trained to use the roads safely and understand traffic safety laws. Hung (2011) further noted that the influence of enforcement on driver behaviour depended on the drivers' attitude towards legislation, whereas education provided a more long-term influence of behavioural change towards safe driving practices and traffic law obedience. Driver behaviour at **signalised junctions** was also found to be an important that factor drivers associated with RTCs. Wang et al. (2002) suggested that suitable side distance, formal and regulated pedestrian activities, and enforcement of appropriate approach speed limits would improve road safety at junctions. Wang et al. (2002) also emphasised that **drivers' perception of safety had significant influence on driving behaviour and task performance**, and it depended on the nature of surrounding information and the drivers' emotional state and personal attributes. Finally, drivers associated road safety to **the type of motorised vehicles**. The likelihood, frequency and severity of road accidents are dependent on the vehicle type and their interaction on multilane roads. It has already been reported that motorcycles were the most involved in road accidents in South-East Asia. This study found "**Road infrastructure and design issues**" as one of the components for Hanoi. Ahmed (2013) confirms that roadway and roadside parameters were significantly linked to road crash occurrence and severity.

4.4. Relationship between driver clusters and components influencing road safety

The relationship between the identified driver clusters and the components of their perception of road safety was investigated using Multinomial Logistic Regression (MLR) and the following results were obtained (Table 7). The correlations between components were also tested (Table 8). The values indicated that components were independent of each other, and therefore, all of them can be used in the MLR. The model fitting information revealed that the final models have the Pseudo R^2 for Jakarta and Hanoi, which is in reasonable range. For studies in arts, humanities and social sciences, the Pseudo R^2 around 0.10 is acceptable (Peterson, 2016).

MLR results in Table 7(a) revealed that relative to Cluster 3 (Teenagers in education riding motorcycles), Cluster 1 (Young employed adults driving cars or motorcycles) in Jakarta perceive that "Type of motorised vehicles" and "Driver behaviour at signalised junctions" have some significant influence on road safety, whereas Cluster 2 (Young and middle age employed males driving cars or motorcycles) only perceive that "type of motorised vehicles" has an influence on road safety. Previous study in Bali (Indonesia) also confirms that signalised junctions are critical site for RTCs by motorists (Dewi, 2021). This is further supported by a study conducted in Malaysia (Ariffin et al., 2010). A recent study from Germany found that accidents in junctions were much more complex than other forms of road accidents, such as straight-road collisions (Sander, 2017). Due to the increasing number of collisions at conflict points, junctions have become a critical place for motorists. As a result, understanding motorist behaviour at junctions needs to be considered to meet motorist expectations in terms of safety when negotiating junctions, particularly in mixed traffic conditions (Ingale et al., 2020).

Additionally, when compared with Cluster 3, Cluster 1 does not realise the importance of "Road safety enforcement and education" as part of improving road safety in Jakarta. In recent years, there have been some road safety education initiatives in Jakarta targeting school children, delivered by the World Health Organisation (WHO, 2015). This means that Cluster 3 (78.5% belong ≤19 age group) are more likely to have some exposure to road safety education and be aware of the value of it in comparison with Cluster 1 (consisting of 91.3% in 20–29 age group); this may be the reason that Cluster 1 does not feel that road safety education is as important as Cluster 3. The study conducted in Yogyakarta (Indonesia) confirms this finding that 48% respondents (a survey of 120 respondents of young population) perceive that current road safety education and campaigns are not strongly affecting road user behaviour (Nurhidayati et al., 2014). The findings revealed that drivers in Jakarta considered type of vehicles as having a significant influence on road safety. This result is similar to the findings from Nurhidayati et al. (2014); 85% of the respondents perceived that type of vehicles have an influence on road safety.

As per the results in Table 7(b) for Hanoi, relative to Cluster 2 (Teenage and young adults in education riding motorcycles), Cluster 1 (Young and middle age employed with postgraduate qualifications driving car) and Cluster 3 (Young adults either employed or in education with undergraduate qualification riding a motorcycle) perceive that "Road safety enforcement and education" and "Road infrastructure design issues" are less likely to have any influence on road safety. Hung and Huyen (2011) initiated some discussion about the benefits of road safety education to reduce accidents. Global Road Safety Partnership has taken some initiative, entitled the "Safety for me, for you and for all" programme, in selected primary schools in Hanoi, Ho Chi Minh City and Danang (GRSP, 2017). This means our reference category (Cluster 2) are more aware of the need for road safety education than the other two clusters. Compared to Cluster 2, Cluster 3 perceives that "Driver behaviour at signalised junctions" has positive associations with road safety, whereas Cluster 2 feels the opposite. Cluster 3 consists of drivers who ride motorcycles while Cluster 2 is composed of car drivers. This may explain their feelings and vulnerability to traffic exposure at junctions. A past study conducted in Bali (Dewi, 2021) also found that those motorists who ignore traffic signals because there are no other cars or motorcycles at the junction as the key concern. Disobedient conduct and reckless driving were also found to be major causes of RTCs in a previous study in Indonesia (Santosa et al., 2017). This result is also in line with a previous study that found that nearly half of all motorcycle accidents were caused by careless or reckless driving (Boni et al., 2018).

5. Conclusions from the study

This study establishes the link between drivers and their perception towards road safety with attention to two major cities in South-East Asia. Five components were identified as influential to drivers' perception of road safety in each city and logistic regression models were developed to examine their level of significance with respect to three sociodemographic cluster groups identified in each city.

The findings revealed that drivers in Jakarta considered type of vehicles and driver behaviour at signalised junctions as significant factors that influenced road safety. Contrary to Decade of Action of Road safety 2011–2020, these drivers did not think that road safety enforcement and education, road infrastructure and roadside facilities, and safe driving practices and behaviour, were important aspects of achieving road safety goals. Hanoi drivers also associated road safety with only two components that included road safety enforcement and education, and driver behaviour at signalised junctions.

It is revealed that even though countries in South-East Asia face similar road safety issues, the magnitude of national mitigation measures varies from country to country. It is also revealed that the level of awareness of road safety among road users in South-East Asia is generally poor. Whereas countries in this region have responded positively towards achievement of the road safety pillars of the decade of action through policy, legislation and strategy, there is considerable gap in their implementation and operationalisation, especially where raising awareness of road safety issues among road users is concerned. Additionally, the study shows that the awareness and perception of road safety among drivers also varies based on various factors like sociodemographic traits. There is a need for road safety managers and stakeholders to understand road user behaviour in order to recommend effective solutions. In a nutshell, the findings reiterate the need for countries to adopt a holistic and collaborative approach to road safety which emphasises safe interaction among road users, vehicles and infrastructure as envisaged in the safe systems.

Conclusively, the components generated in our PCA analysis clearly show that the perception of road users towards road safety revolves around the three road crash causation factors of infrastructure, vehicles and humans, as suggested by wider literature. Therefore, this study confirms that road safety improvement strategies and policies should adopt an integrated approach that facilitates secure interaction of infrastructure, vehicle and human factors since road user behaviour is dependent on all these factors.

CRediT authorship contribution statement

Moris Thibenda: Writing – original draft, Validation, Formal analysis, Data curation. **Dewa Made Priyantha Wedagama:** Writing – review & editing. **Dilum Dissanayake:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Driver Cluster Classification

See Tables A1 and A2.

Table A1
Driver Cluster Classification with Most Frequent Categories – Jakarta.

Main Category	Subcategories	Clusters		
		1 (57.5%)	**2 (22.9%)**	**3 (19.7%)**
Gender	Male	55.4%	90.2%	57.0%
	Female	44.6%	9.8%	43.0%
Age	≤19	5.1%	3.2%	78.5%
	20–29	91.3%	32.6%	21.5%
	30–39	3.6%	39.1%	0.0%
	40–49	0.0%	16.3%	0.0%
	50 ≥	0.0%	8.8%	0.0%
Occupation	Public Employee	2.7%	16.1%	0.0%
	Private Employment	59.7%	56.5%	0.0%
	Self Employed	5.2%	18.3%	5.1%
	Housewife/husband	0.0%	0.0%	0.0%
	Unemployed	2.4%	0.0%	0.0%
	Retired	5.1%	5.1%	0.0%
	Student	25.0%	4.0%	94.9%
1 vehicle in household	No	77.9%	51.1%	67.1%
	Yes	22.1%	48.9%	32.9%
Driving experience	<1 year	7.4%	35.9%	4.9%
	1–5 year(s)	29.8%	18.0%	31.1%
	5–10 years	37.7%	32.6%	51.9%
	10–20 years	14.9%	10.3%	12.1%
	>20 years	10.2%	3.2%	0.0%
Driving frequency	Everyday	63.6%	54.6%	92.4%
	2 or 3 times a week	21.2%	28.7%	7.6%
	once a week	14.1%	8.4%	0.0%
	once a month	1.1%	8.2%	0.0%
	hardly ever	0.0%	0.0%	0.0%
Vehicle type	Motorcycle	50.2%	51.1%	93.7%
	Sedan/SUV	43.7%	45.7%	6.3%
	Bus/Minibus	6.1%	3.2%	0.0%
	Others	0.0%	0.0%	0.0%

Table A2
Driver Cluster Classification with Most Frequent Categories – Hanoi.

Main Category	Subcategories	Clusters		
		1 (20.4%)	**2 (23.5%)**	**3 (56.1%)**
Age	≤19	8.8%	53.3%	3.1%
	20–29	17.3%	40.7%	83.6%
	30–39	34.4%	6.0%	10.2%
	40–49	37.5%	0.0%	3.1%
	50 ≥	2.0%	0.0%	0.0%
Occupation	Public Employee	25.5%	0.0%	10.3%
	Private Employment	22.3%	6.1%	42.3%
	Self Employed	35.0%	5.9%	8.0%
	Housewife/husband	0.0%	0.0%	0.0%
	Unemployed	0.0%	0.0%	0.0%
	Retired	0.0%	0.0%	0.0%
	Student	17.2%	88.0%	39.4%
Income	<5 million VND	0.0%	90.2%	33.2%
	5 to 10 million VND	8.7%	9.8%	32.3%

(*continued on next page*)

Table A2 (*continued*)

Main Category	Subcategories	Clusters		
		1 (20.4%)	**2 (23.5%)**	**3 (56.1%)**
	10 to 15 million VND	21.3%	0.0%	25.6%
	15 to 30 million VND	35.0%	0.0%	8.9%
	30 to 50 million VND	35.0%	0.0%	0.0%
	>50 million VND	0.0%	0.0%	0.0%
Qualification	High school	9.0%	75.0%	25.5%
	Vocational school	10.7%	0.0%	0.0%
	Undergraduate	27.8%	25.0%	66.4%
	Postgraduate	52.5%	0.0%	8.1%
	Other	0.0%	0.0%	0.0%
No vehicle in household	No	98.8%	88.3%	99.1%
	Yes	1.2%	11.7%	0.9%
2 vehicles in household	No	63.7%	87.0%	95.5%
	Yes	36.3%	13.0%	4.5%
Vehicle type	Motorbike	28.8%	71.7%	96.4%
	Car	71.2%	28.3%	3.6%
	Bus	0.0%	0.0%	0.0%
Driving experience	<1 year	15.0%	98.9%	14.6%
	1–5 year(s)	37.5%	1.1%	52.7%
	5–10 years	27.5%	0.0%	32.7%
	10–20 years	20.0%	0.0%	0.0%
	>20 years	0.0%	0.0%	0.0%
Driving frequency	Everyday	70.0%	90.1%	89.5%
	2 or 3 times a week	30.0%	9.9%	11.5%
	once a week	0.0%	0.0%	0.0%
	once a month	0.0%	0.0%	0.0%
	hardly ever	0.0%	0.0%	0.0%

References

Ahmed, I., 2013. Road infrastructure and road safety. Transport Commun. Bull. Asia Pacific 83, 19–25.

Akgün, N., Dissanayake, D., Thorpe, N., Bell, M.C., 2018. Cyclist casualty severity at roundabouts–To what extent do the geometric characteristics of roundabouts play a part? J. Saf. Res. 67, 83–91.

Akgün, N., Daniels, S., Bell, M.C., Nuyttens, N., Thorpe, N., Dissanayake, D., 2021. Exploring regional differences in cyclist safety at roundabouts: A comparative study between the UK (based on Northumbria data) and Belgium. Accid. Anal. Prev. 150, 105902.

Akgün-Tanbay, N., Campisi, T., Tanbay, T., Tesoriere, G., Dissanayake, D., Lee, J., 2022. Modelling Road User Perceptions Towards Safety, Comfort and Chaos at Shared Space: The Via Maqueda Case Study, Italy. J. Adv. Transport. Article ID: 4979496.

Ali, F., Dissanayake, D., Bell, M., Farrow, M., 2018. Investigating car users' attitudes to climate change using multiple correspondence analysis. J. Transp. Geogr. 72, 237–247.

Ariffin, A.H., Jawi, Z.M., Isa, M.H., Kassim, K.A.A., Voon. W.S., 2010. Pedestrian Casualties in Road Accidents – Malaysia Perspective. Paper presented at 1st MIROS Road Safety Conference (ROSCON), Kuala Lumpur, Malaysia.

Arnau-Sabatés, L., Sala-Roca, J., Jariot-Garcia, M., 2012. Emotional abilities as predictors of risky driving behaviour among a cohort of middle-aged drivers. Accid. Anal. Prev. 45, 818–825.

Atombo, C., Wu, C., Zhong, M., Zhang, H., 2016. Investigating the motivational factors influencing drivers' intentions to unsafe driving behaviours: Speeding and overtaking violations. Transport. Res. (part F) 43, 104–121.

Bekasi Statistics, 2022 Badan Pusat Statistik (bps.go.id).

Boni, J.A., Chowdhury, T.R., Das, S.S., 2018. A Study on causes of road accidents at Dhaka to Comilla Highway, *Asian Journal of Innovative Research in Science*. Eng. Technol. 1 (9), 6–13.

Borowsky, A., Oron-Gilad, T., 2013. Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. Accid. Anal. Prev. 59, 548–565.

Borrell, C., Plasencia, A., Huisman, M., Costa, G., Kunst, A., Andersen, O., Bopp, M., Borgan, J.K., Bogor Statistics, 2022. Bogor City Statistics Central Agency (bps.go.id).

Bradburn, K., 2022. The Bonferroni Correction Method Explained, toptipbio.com.

Brown, J.D., (2009). JALT Testing and Evaluation SIG Newsletter. 13 (3) November 2009 (p. 20–25), Microsoft Word - 5-Brown Article.doc (jalt.org).

Bui, H.T., Saadi, I., Cools, M., 2020. Investigating on-road crash risk and traffic offences in Vietnam using the motorcycle rider behaviour questionnaire (MRBQ). Saf. Sci. 130, 104868.

Chekijian, S.A., 2012. Practices, Attitudes and Perceptions towards Road Safety in Yerevan, Republic of Armenia. Assoc. Adv. Automotive Med. 56, 191–200.

Chen, G., 2010. Road traffic safety in African countries – status, trend, contributing factors, countermeasures and challenges. Int. J. Injury Control Safety Promotion 17 (4), 247–255. https://doi.org/10.1080/17457300.2010.490920.

Cordellieri, P., Baralla, F., Ferlazzo, F., Sgalla, R., Piccardi, L., Giannini, A.M., 2016. Gender effects in young road users on road safety attitudes, behaviors and risk perception. Front. Psychol. 7, 1412.

Costa, P.S., Santos, N.C., Cunha, P., Cotter, J., Sousa, N., 2013. The Use of Multiple Correspondence Analysis to Explore Associations between Categories of Qualitative Variables in Healthy Ageing. J. Aging Res. 2013, 1–12.

Das, S., Sun, X., 2016. Association knowledge for fatal run-off-road crashes by multiple correspondence analysis. IATSS Research 39 (2), 146–155.

Depok Statistics, 2022. Depok City Central Statistics Agency (bps.go.id).

Deffenbacher, J.L., Oetting, E.R., Lynch, R.S., 1994. Development of a driving anger scale. Psychol. Rep. 74 (1), 83–91.

Dewi, A.A.D.P., 2021. Investigating Motorists Perceptions towards Road Safety. Civil Eng. Arch. 9 (5), 1339–1346.

Dissanayake, S., 2004. Comparison of severity affecting factors between young and older drivers involved in single vehicle crashes. IATSS research 28 (2), 48–54.

Edquist, J., Rudin-Brown, C., Lenne, M.G., 2009. Road design factors and their interactions with speed and speed limits. Monash University Accident Research Centre 30, 24.

Field, A., 2017. Discovering statistics using IBM SPSS statistics. North American edition. sage.

Fruhen, L.S., Flin, R., 2015. Car driver attitudes, perceptions of social norms and aggressive driving behaviour towards cyclists. Accid. Anal. Prev. 83, 162–170.

General Statistics Office in Vietnam, 2019. Completed Results of the 2019 Vietnam Population and Housing Censes, Statistical Publishing House (Vietnam), ISBN978-604-75-1532-5.

Yannis, G., Kanellopoulou, A., Aggeloussi, K., Tsamboulas, D., 2005. Modelling driver choices towards accident risk reduction. Saf. Sci. 43 (2005), 173–186.

GRSP, 2017. Road safety for school children, a focus for Total Vietnam | | GRSP | Global Road Safety Partnership (grsproadsafety.org).

Greenacre, M., Blasius, J. (Eds.), 2006. Multiple correspondence analysis and related *methods. CRC Press.*

Gubry, P., Huong, L.T., Thieng, N.T., 2009. Disparities in the city: Poverty and urban environment in Hanoi and Ho Chi Minh City (Vietnam). 26th IUSSP International Population Conference Marrakesh, 27th September-2nd October 2009.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L., 1998. Multivariate data analysis (Vol. 5, No. 3, pp. 207-219). Upper Saddle River, NJ: Prentice hall.

Hair, J.F., Gabriel, M., Patel, V., 2014. AMOS covariance-based structural equation modeling (CB-SEM): Guidelines on its application as a marketing research tool. Brazilian J. Market. 13 (2).
Hazen, A., Ehiri, J.E., 2006. Road traffic injuries: hidden epidemic in less developed countries. J. Natl Med. Assoc. 98 (1), 73.
Hung, V.K., Huyen, T., 2011. Education influence in traffic safety: A case study in Vietnam. IATSS Research 34 (2), 87–93.
Indonesia Police Office, https://jakarta.bps.go.id/indicator/17/786/1/jumlah-kendaraan-bermotor-menurut-jenis-kendaraan-unit-di-provinsi-dki-jakarta.html, accessed on Feb 2022.
Ingale, A., Sahu, P., Bajpai, R., Maji, A., Sarkar, A., 2020. Understanding driver behaviour at intersection for mixed traffic conditions using questionnaire survey. In: Mathew, T., Joshi, G., Velaga, N., Arkatkar, S. (Eds.), Transportation Research. Lecture Notes in Civil Engineering, Springer, Singapore, p. 45.
Jakarta Statistics, 2022. BPS Provinsi DKI Jakarta.
Janssen, K.M.J., Mulder, P., Yudhistira, M.H., 2021. Spatial sorting of rich versus poor people in Jakarta. Bull. Indonesian Economic Stud. https://doi.org/10.1080/00074918.2021.1876209.
Jones, S., Odero, K., Adanu, M.E., 2020. Road crashes in Namibia: Challenges and opportunities for sustainable development. Develop. Southern Africa 37 (2), 295–311.
Johnson, A., Wichern, D.W., 2007. Applied Multivariate Correspondence Analysis, Prentice-Hall, Upper Saddle River, NJ, USA, 6th edition.
Juan de Oña, Rocio de Oña, Laura Eboli, Carmen Forciniti and Gabriella Mazzulla, 2014. How to identify the key factors that affect driver perception of accident risk. A comparison between Italian and Spanish driver behaviour. Accident Analysis & Prevention 73, 225–235.
Kayani, A., King, M., Fleiter, J., 2011. Fatalism and road safety in developing countries, with a focus on Pakistan. J. Australasian College Road Safety 22 (2), 41–47.
Kitamura, Y., Hayashi, M., Yagi, E., 2018. Traffic problems in Southeast Asia featuring the case of Cambodia's traffic accidents involving motorcycles. IATSS Research 42 (4), 163–170.
Lajunen, T., Summala, H., 1995. Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. Personality Individ. Differ. 19 (3), 307–318.
Lee, C., 2016. To Uberize or Not to Uberize? Opportunities and Challenges in Southeast Asia's Sharing Economy. ISEAS 33, 2016.
Li, Q., Tay, R., 2013. Improving drivers' knowledge of road rules using digital games. Accid. Anal. Prev. 65, 8–10.
McIlvenny, S., Mahrouqi, F.A., Busaidi, T.A., 2004. Rear seat belt use as an indicator of safe road behaviour in a rapidly developing country. JRSH 124 (6), 280-283.
Norusis, M.J., 2011. IBM SPSS statistics 19 procedures companion. Reading, MA, Addison-Wesley.
Nurhidayati, Z.A., Petros, S., Malkhamah, S.S., 2014. Road Safety Perception towards Sustainable Transport: A Case of road Safety Programs in Yogyakarta, *17th FSTPT International Symposium*, Jember University, 22-24 August 2014.
Ortúzar, J.d.D., Willumsen, L.G. (Eds.), 2011. Modelling Transport. Wiley.
Peterson, O., 2016. Re: What is the acceptable r-squared value?. Retrieved from: https://www.researchgate.net/post/what_is_the_acceptable_rsquared_value/57cfcb0d3d7f4be2fb47f515/citation/download.
Papadimitriou, E., Theofilatos, A., Yannis, G., 2013. Patterns of pedestrian attitudes, perceptions and behaviour in Europe. Saf. Sci. 53, 114–122.
Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K., 1990. Errors and violations on the roads: a real distinction? Ergonomics 33 (10–11), 1315–1332.
Ribbens, S., Dewulf, J., Koenen, F., Mintiens, K., De Sadeleer, L., de Kruif, A., Maes, D., 2008. A survey on biosecurity and management practices in Belgian pig herds. Preventive Veterinary Med. 83 (3-4), 228–241.
Sam, E.F., Brijs, K., Daniels, S., Brijs, T., Wets, G., 2019. Construction and validation of a public bus passenger safety scale. Transport. Res. (Part F) 66, 47–62.
Sami, A., Najafi, A., Yamini, N., Moafian, G., Aghabeigi, M.R., Lankarani, K.B., Heydari, S.T., 2013. Educational level and age as contributing factors to road traffic accidents. Chinese J. Traumatol. 16 (5), 281–285.
Sayed, T., Abdelwahab, W., Navin, F., 1995. Identifying accident-prone locations using fuzzy pattern recognition. J. Transp. Eng. 121 (4), 352–358.
Schlögl, M., 2020. A multivariate analysis of environmental effects on road accident occurrence using a balanced bagging approach. Accid. Anal. Prev. 136, 105398.
Speck, J., 2012. Walkable City: How Downtown Can Save America, One Step at A Time. Farrar, Straus and Giroux, New York, USA.
Sander, U., Opportunities and limitations for intersection collision intervention (2017) a study of real world left turn across path' accidents, Accident Analysis and Prevention, 99, 342–355.
Santosa, S.P., Mahyuddin, A.I., Sunoto, F.G., 2017. Anatomy of injury severity and fatality in Indonesian traffic accidents. J. Eng. Technol. Sci. 49 (3), 412–422.
Senbil, M., Zhang, J., Fujiwara, A., 2007. Motorization in Asia: 14 countries and three metropolitan areas. IATSS research 31 (1), 46–58.
Şimşekoğlu, Ö., Nordfjærn, T., Zavareh, M.F., Hezaveh, A.M., Mamdoohi, A.R., Rundmo, T., 2013. Risk perceptions, fatalism and driver behaviors in Turkey and Iran. Saf. Sci. 59, 187–192.
Sivasankaran, S.K., Balasubramanian, V., 2020. Investigation of pedestrian crashes using multiple correspondence analysis in India. Int. J. Injury Control Safety Promot. 27 (2), 144–155.
Soehodho, S., 2009. Road accidents in Indonesia. IATSS Research 33 (2), 122–124.
Tangerang Statistics, 2022. Badan Pusat Statistik (bps.go.id).
Theofilatos, A., Yannis, G., 2014. A review of the effect of traffic and weather characteristics on road safety. Accid. Anal. Prev. 72, 244–256.
Tuan, V.A., 2015. Mode choice behavior and modal shift to public transport in developing countries-the case of Hanoi city. J. Eastern Asia Soc. Transport. Stud. 11, 473–487.
Turner, S., 2020. Informal motorbike taxi drivers and mobility injustice on Hanoi's streets. Negotiating the curve of a new narrative. J. Transp. Geogr. 85, 102728.
Ulleberg, P., Rundmo, T., 2003. Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. Saf. Sci. 41 (5), 427–443.
Vanlaar, W., Yannis, G., 2006. Perception of road accident causes. Accid. Anal. Prev. 38 (1), 155–161.
Wang, B., Hensher, D.A., Ton, T., 2002. Safety in the road environment: a driver behavioural response perspective. Transportation 29 (3), 253–270.
Wedagama, D.M.P., Bennett, S., Dissanayake, D., 2020. Analyzing Pedestrian Perceptions towards Traffic Safety Using Discrete Choice Models. International Journal on Advanced Science, Engineering and Information Technology 2020, 10(6), 2394-2401.
Wedagama, D.M.P., Dissanayake, D., 2020. A model of latent class multinomial logit to investigate motorcycle accident injuries. Eng. Appl. Sci. Res. 47 (4), 422–429.
Wedagama, D.M.P., Dissanayake, D., 2009. Analysing motorcycle injuries on arterial roads in Bali using a Multinomial Logit model 7, 363.
Wedagama, D.M.P., Roger, B., Dissanayake, D., 2008. The influence of urban land use on pedestrians casualties: case study area: Newcastle upon Tyne, UK. IATSS research 32 (1), 62–73.
Wedagama, D.M.P., Dissanayake, D., Bird, R., 2011. Modelling the relationship between male and female pedestrian accidents and land use characteristics (case study: Newcastle upon Tyne, UK). In: Proceedings of the Eastern Asia Society for Transportation Studies, vol. 8, pp. 345-345.
Widyastuti, H., Mulley, C., Dissanayake, D., 2007. Binary Choices Model to Value Motorcyclist's Slight Injury Cost in Surabaya, Proceedings of the Eastern Asia Society for Transportation Studies.
Wright, R.E., 1995. Logistic regression. In: Grimm, L.G., Yarnold, P.R. (Eds.), Reading and understanding multivariate statistics. American Psychological Association, pp. 217–244.
World Bank, 2021. Country Classification World Bank Country and Lending Groups – World Bank Data Help Desk.
World Health Organization (WHO), 2018a. Global status report on road safety 2018, www.who.int/publications/i/item/9789241565684.
World Health Organization (WHO), 2018b. Road Traffic Injuries. Road traffic injuries (who.int).
World Health Organization (WHO), 2015. United Nations Road Safety Collaboration, https://www.who.int/roadsafety/en.
Yoshida, H., Tsukatsune, K., Niida, S., 2020. Analysis of Mental Model of Users with Network Malfunction. In: Yamamoto, S., Mori, H. (Eds.), Human Interface and the Management of Information. Designing Information. HCII 2020. Lecture Notes in Computer Science, 12184. Springer, Cham. https://doi.org/10.1007/978-3-030-50020-7_22.
Zafar, S.N., Canner, J.K., Nagarajan, N., Kushner, A.L., Gupta, S., Tran, T.M., Groen, R.S., 2018. Road traffic injuries: cross-sectional cluster randomized countrywide population data from Four low-income countries. Int. J. Surg. 52, 237–242.