The increase in population is accompanied by an increase in the number of vehicles. It is inevitable that the number of vehicle accidents will also increase, which can be caused by various factors. Driver factors reviewed in this study include socioeconomic characteristics, movement characteristics, accident characteristics, and driver behavior characteristics. the purpose of this study is to study the vehicle accident model using interviews and Driving Behavior questionnaires with a total of 307 motorist respondents who have experienced accidents. Driver factors reviewed in this study include socioeconomic characteristics, movement characteristics, accident characteristics, and driver behavior characteristics using interviews and Driving Behavior questionnaires with a total of 307 motorist respondents who have experienced accidents.

This investigate used SEM (Structural Equation Modeling) with SmartPLS computer software. Two-wheeled vehicle accident modeling results $Y=-0.234X1+0.153X3+ +ei2$; $R^2=0.102$. The greatest influence occurs in the characteristics of driver behavior (X3), namely Ordinary Violation, and for four-wheeled vehicle accident modeling results, $Y=-0.343X1+0.284X3+ei2$; $R^2=0.217$. The greatest influence occurs in driver behavior characteristics (X3), namely Ordinary Violation. Ordinary Violation is defined as a deliberate deviation from the rule of law.

Thus, from the research results, the most influential variable was the behavior of drivers who committed ordinary violations such as ignoring speed limits, breaking through intersections, and driving under the influence of alcohol. So, there needs to be collaboration between the police and related parties in tackling accidents and reducing the risk of traffic accidents, such as long as socialization or information through newspapers or electronic media to the public in Jayapura City regarding the importance of collective awareness of driving safety

***Keywords:** vehicle accidents, driving behavior, structural equation modeling, traffic accidents, motor vehicle drivers, car driver behavior, driving characteristics*

UDC 656

DOI: 10.15587/1729-4061.2024.306781

IDENTIFYING THE VEHICLE ACCIDENT MODELS BASED ON DRIVING BEHAVIOR FACTORS USING STRUCTURAL EQUATION MODELING

Fadila Ardi Putri Damayanti
Corresponding author
Master's Student in Civil Engineering
Department of Civil Engineering*
E-mail: fadiladamayanti@student.ub.ac.id

Muhammad Zainul Arifin
Doctor of Civil Engineering
Department of Civil Engineering*

Fauzul Rizal Sutikno
Doctor of Urban Regional and Planning
Departmen Urban Regional and Planning*

Muh Miftahulkhair
Master of Civil Engineering
Department of Civil Engineering
Universitas Sulawesi Barat
Prof. Dr Baharuddin Lopa str., SH,
Majene, Indonesia, 91412
*Brawijaya University
Veteran str., Malang, East Java, Indonesia, 65145

Received date 12.04.2024
Accepted date 17.06.2024
Published date 28.06.2024

***How to Cite:** Damayanti, F. A. P., Arifin, M. Z., Sutikno, F. R., Miftahulkhair, M. (2024). Identifying the vehicle accident models based on driving behavior factors using structural equation modeling. Eastern-European Journal of Enterprise Technologies, 3 (3 (129)), 85–93. https://doi.org/10.15587/1729-4061.2024.306781*

1. Introduction

Traffic accidents result from the malfunctioning of a complex system involving vehicles, road infrastructure, road users, and their communications [1]. They stand as the principal cause of death globally and are projected to rank fifth by 2020 [2] with 91 % of fatalities occurring on roads [3]. Previous studies have identified human factors as the primary contributors to accidents [4] Research indicates a positive and significant correlation between driving behavior and crash involvement, with driving behavior posing a 50 % increased risk of accidents [5, 6]. The prediction model obtained is $Y=-0.203X1+(-0.179X2)+0.214X3+0.536X4$. The highest significant path coefficient of 0.536 is found in the driver behavior variable ($X4$) with the highest factor weight of 0.638, namely driving behavior. with the highest factor weight of 0.638, namely driving behavior under the influence of alcoholic beverages ($X4.7$). Accident modeling results $Y=0.299X1+0.154X2+0.077X3+0.554X4$. The first largest influence on the probability of a crash is the driving behavior characteristic ($X4$) exceeding the speed ($X4.10$). The more often a driver exceeds the rate, the higher the chance of a crash. The second most significant influence of socio-economic characteristics ($X1$) is the age indicator ($X1.2$), the more mobility in productive age, the higher the risk of accidents [7]. Therefore, it is necessary to add billboards as a warning to reduce vehicle speed, provide shock markers or don't forget to rest when traveling long distances [7, 8].

In Indonesia, the National Police reported a staggering 116,441 fatalities from accidents in 2019, emphasizing the prevalence of the issue. The research variable indicators used are socio-economic characteristics, driving equipment and preparation, driver habits, driver behavioral characteristics, accident characteristics [7] and the effects of overloading on heavy vehicles [9]. Despite previous studies focusing on human factors, they often employ outdated analysis methods and lack detailed insights into the reasons of accidents and their relationship with human factors.

Therefore, studies that are devoted at predicting the development of motorcycle accidents are scientifically relevant and necessary to minimize the incidence of accidents.

2. Literature review and problem statement

Research conducted by [7] found that the most influential factor is driver behavior. The highest weight is driving under the influence of alcohol and the second highest weight is driving in a drowsy state.

Research conducted by [8] shows that age is a significant factor in accidents. However, from the model, it is found that the most influential weight factor is driver behavior but it is not explained what kind of driver behavior affects.

However, these two studies only analyzed driver behavior in general, so they are not updated and cannot describe driver behavior as a cause of accidents in detail. Therefore, in this study, a driving behavior questionnaire (DBQ) is used as a guide to determine the driver behavior that causes the high number of accidents.

Despite advancements in transportation, road accidents persist as a pressing challenge. They pose significant risks to road users, resulting in both physical and material losses, necessitating efforts for prevention to ensure transportation safety. Despite initiatives to report these issues, such as enhancements in road conditions and lighting, research suggests that these improvements can impact drivers' speed perceptions, influencing road safety [1].

Research conducted by [10, 11] shows that age is a contributing element to the occurrence of accidents, but this research does not explain how age is related to its impact on the occurrence of accidents. However, this study used a random sample, where respondents were drawn from those who had or had not experienced an accident.

Various factors, including individuals, roads, vehicles, and environmental/weather conditions, contribute to road accidents, with drivers' perceptions and behaviors playing a crucial role. Previous studies have investigated driving rule violations and aggressive behaviors concerning traffic accidents to comprehend driver conduct and performance. Some researchers have assessed the influence of these factors on driver behavior and attributed road accidents to infrastructure degradation and driver inexperience [12]. These studies focus on risk analysis using objective data such as crash counts, infrastructure conditions, and policies, but do not go in-depth into driver behavioral factors or social impacts that affect crash risk. Factors such as cell phone use while driving, compliance with traffic laws, and aggressive driving behavior may not be explored in depth.

Research conducted by [13] has highlighted that exceeding speed is the most significant driving behavior contributing to the probability of a crash. Over-speeding during overtaking maneuvers can lead to collisions resulting in fatal injuries. Several factors contribute to this, including non-compliance with traffic regulations and exceeding speed limits. While research has quantified the contribution of speed to crashes, aspects such as the influence of rider behavior, experience and ability to control the motorcycle in emergency conditions may not have been fully explored. For example, how daily driving habits and compliance with traffic laws affect crash risk. Research may have a narrow focus due to the study objective of measuring the direct effect of one particular variable (such as speed) on crashes.

Disobeying traffic signs and driving above speed limits are among the primary factors contributing to motor vehicle accidents [12] Socioeconomic status can be approximated by factors such as education level, employment status, and neighborhood income [14] The study focused on driving behavior and its influence on crashes, but did not deeply explore how social norms and driving culture in Saudi Arabia influence driving behavior. This includes the influence of social norms on the acceptability of aggressive or reckless driving behavior. Resource constraints and priorities in research may also influence which aspects of the problem are considered important to explore. Research may be focused on the variables deemed most significant or easiest to measure.

Lower socioeconomic status is associated with a higher risk of traffic injuries [15] and fatalities [16]. While the analysis focuses on the direct causes of crashes such as traffic violations and crash severity, indirect causes such as socio-economic influences or demographic factors may not have been studied in depth. The tendency to focus on easily accessible data such as accident reports may limit the ability to assess less well-documented indirect factors or long-term impacts.

Additionally, demographic factors such as education level, gender, age, and occupation have been found to influence the number of traffic fatalities in previous studies [17]. This study focuses only on male drivers, reflecting the unique context of Saudi Arabia where, traditionally, women were not allowed to drive until recently. Women drivers' driving behaviors, risk perceptions, and responses to traffic safety interventions remain largely unexplored in this context.

All this allows to assert that it is prudent to conduct studies on accidents. how this problem of accidents has not been able to reduce the number of occurrences. This study aims to identify factors that contribute to accidents, focusing on driving behavior and socio-economic characteristics of drivers, using Structural Equation Modeling (SEM) to ensure safe and secure transportation. Despite this, vehicle accidents remain a significant challenge, posing a risk of physical and material losses, which highlights the urgent need for accident prevention to ensure transportation safety [8]. the expected result is to get influential factors in the occurrence of accidents. So that it can find solutions to reduce the incidence of accidents that occur in the future.

3. The aim and objectives of the study

This study aims to ascertain the characteristics of motor vehicle drivers. This will provide new insights for practitioners and scientists regarding these characteristics and to find factors that cause accidents help reduce accidents.

To achieve this aim, the following objectives are accomplished:

– to determining the most critical factors among the parameters that influence accidents: socioeconomic characteristics ($X1$), travel pattern characteristics ($X2$), behavioral characteristics ($X3$) to crash characteristics (Y);

– to develop a model predicting the likelihood of motor vehicle accidents in Jayapura.

4. Materials and methods

This research was conducted to analyze what are the accident factors that affect accidents. The sampling technique was carried out using the Isaac and Michael formula [8] with a sample size of 307 people based on the number of accidents in Jayapura City. Sampling in this study used purposive sampling, where respondents were selected only who had experienced an accident so that the results of the analysis could be accurately known about the factors that caused the accident. The questionnaire was addressed to drivers of two-wheeled and four-wheeled vehicles. So that the modeling results will be significant and valid. The research method used is interview and questionnaire techniques using SmartPLS SEM (Structural Equation Modeling) software to find the accident model.

Previous research has predominantly relied on regression analysis, employing a relatively simple method. Thus, this study aims to complement existing literature by delving deeper into the analysis of driving behavior and socioeconomic characteristics associated with the high incidence of accidents on Jayapura City roads, employing SEM (Structural Equation Modeling) with SmartPLS software.

SEM is particularly relevant to this study due to its ability to analyze relationships among multiple variables. Initially developed by Sewall Wright in 1934, SEM, originally known as path analysis, has evolved into a comprehensive analytical tool [18]. SEM enables the examination of relationships between latent constructs and their indicators, as well as interactions among latent constructs and direct measurement errors [19]. Notably, SEM allows for the simultaneous analysis of multiple dependent and independent variables [20]. Given its suitability to the research objectives, SEM was chosen as the analytical method for this study. The research variables to be analyzed can be seen in Table 1 below.

Table 1 shows the research design based on each indicator. socioeconomic characteristics ($X1$): age, gender, education, jobs and revenue. movement characteristics ($X2$): distance traveled and travel time. driving behavior ($X3$): Error (ignoring the speed limit, not looking at the rearview mirror, failing to turn on the turn signal, trying to overtake from the left side, almost crashing, not noticing pedestrians crossing, braking suddenly), Lapse (entering the wrong lane, forgetting to park the vehicle, realizing the mistake but still doing it, misreading the signs, having difficulty remembering the road clearly, bumping into other drivers), ordinary violation (ignoring the speed limit, running traffic lights, overtaking from the left, driving under the influence of alcohol, not maintaining a safe distance) aggressive violation (honking to show anger, forcing through a closed lane, speaking harshly on the road, chasing other drivers, racing at red lights). In the research design in Table 1, all questions were derived from previous research. These questions or variables certainly affect the chance of an accident occurring.

The illustration of the flow chart of the data collection and analysis process is the steps taken by the author in obtaining data and conducting data analysis. The first thing that must be done is an interview of respondents who have experienced an accident using a questionnaire and interviews related to socio-economic characteristics, movement, behavior and probability of accidents. then reduce and compile into ordinal, nominal and ratio scales. After that, analysis was carried out using SEM PLS inner model and outer model and obtained the results of predicting the probability of an accident, to determine the characteristics of vehicle and accident characteristics, as shown in Fig. 1.

The flowchart illustrates the data collection and analysis process, detailing the steps involved in acquiring and processing the data for analysis.

Table 1

Survey Design

Notation	Question	Measurement Scale
Socioeconomic characteristics ($X1$)		
$X1.1$	Age	Interval
$X1.2$	Gender	Nominal
$X1.3$	Education	Ordinal
$X1.4$	Jobs	Nominal
$X1.5$	Revenue	Interval
Movement Characteristics ($X2$)		
$X2.1$	Distance traveled	Interval
$X2.2$	Travel time	Interval
Driving Behavior ($X3$)		
$X3.1$	Error	Ordinal
$X3.2$	Lapse	Ordinal
$X3.3$	Ordinary Violation	Ordinal
$X3.4$	Aggressive Violation	Ordinal
Chance of Accident (Y)		
$Y1$	Frequency of having an accident	Ordinal
$Y2$	Frequency of near-accidents	Ordinal
$Y3$	Type of Accident	Ordinal
$Y4$	Causes of Accidents	Ordinal
$Y5$	Injury	Ordinal
$Y6$	Accident Time	Ordinal
$Y7$	Collision Type	Ordinal

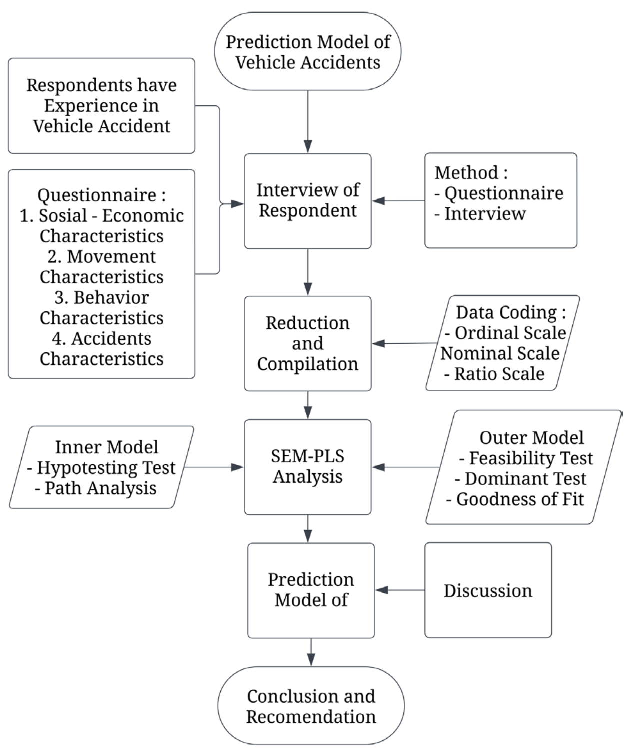

Fig. 1. Flowchart

5. Results of probability model of vehicle accident

5. 1. The most critical factor among the parameters affecting accidents

To find out which variables and indicators are most influential in affecting accidents, it is necessary to conduct the following analysis:

5. 1. 1. Two wheeled vehicle feasibility/validity test

This test model aims to describe how well the indicators in this study can be used as instruments for measuring latent variables. With a significance weight <0.05 (5 %). VIF value <10.

Table 2 presents the outer model measurement results for formative indicators. Variable X1 (Socio-economic characteristics) exhibits all indicators with a VIF smaller than five and valid items with a p-value smaller than 0.05. However, only two out of five items in variable *X*1 are deemed valid, leading to the elimination of items *X*1.3, *X*1.4, and *X*1.5.

Table 2

Outer model measurement of formative indicators

Dimensions	Item	VIF	P-values Outer weight	Description
*X*1 (Socio-economic characteristics)	*X*1.1	1.793	0.016	Valid
	*X*1.2	1.055	0.003	Valid
	*X*1.3	1.231	0.899	Invalid
	*X*1.4	1.597	0.304	Invalid
	*X*1.5	2.490	0.208	Invalid
*X*2 (Movement Characteristics)	*X*2.1	1.370	0.180	Invalid
	*X*2.2	1.370	0.848	Invalid
Y (Chance of Accident)	*Y*1	1.338	0.000	Valid
	*Y*2	1.303	0.000	Valid
	*Y*3	1.220	0.977	Invalid
	*Y*4	1.045	0.504	Invalid
	*Y*5	1.081	0.386	Invalid
	*Y*6	1.044	0.234	Invalid
	*Y*7	1.186	0.716	Invalid

Variable *X*2 (movement characteristics) displays all indicators with a VIF smaller than five and valid items with a p-value smaller than 0.05. Unfortunately, there are no valid items out of the two in variable *X*2, necessitating its elimination from this model.

Reflective indicators (variable X3).

Table 3 present of the tests to measure the Outer model with reflective indicators include assessments of convergent validity, discriminant validity, composite reliability, average variance extracted (AVE), and Cronbach's Alpha.

Table 3

Table of outer model measurement results reflective indicators (after modification)

Latent variable	Item	Outer loading	Description	Cronbach's Alpha	Composite reliability	AVE
*X*3 (driving behavior characteristics)	*X*3.1	0.784	Valid	0.856	0.900	0.693
	*X*3.2	0.819	Valid			
	*X*3.3	0.903	Valid			
	*X*3.4	0.820	Valid			

According to the table above, it is evident that all indicators exhibit an outer loading value exceeding 0.700 (Valid), indicating the validity of all indicators. Within each variable, certain items predominantly reflect these variables. Specifically, in variable *X*3 (driver behavior characteristics), the most dominant indicator is *X*3.3, representing Ordinary Violation behavior, with the highest loading factor of 0.903.

Referring to the table 3, it is evident that all indicators exhibit construct-forming cross-loading values that are higher within their respective variables compared to the loading values on other variables. Specifically, the loading factor value of indicators within variable *X*3 remains higher than those of other variables.

The Cronbach's Alpha value for variable *X*3 is 0.856, exceeding the threshold of 0.7, indicating that variable *X*3 (driver behavior characteristics) is reliable. Additionally, the composite reliability figure of 0.900 surpasses 0.7, categorizing variable *X*3 (driver behavior characteristics) as highly reliable. Discriminant validity, as assessed by the average variance extracted (AVE), confirms the validity of variable *X*3, with a value of 0.693 exceeding 0.5.

5.1.2. Four wheeled vehicle feasibility/validity test

Table 4 present the outer model measurement results from formative indicators.

Table 4

Outer model measurement results from formative indicators

Dimensions	Item	VIF	P-values Outer weight	Description
*X*1 (socio-economic characteristics)	*X*1.1	1.141	0.906	Invalid
	*X*1.2	1.140	0.353	Invalid
	*X*1.3	1.246	0.021	Valid
	*X*1.4	1.253	0.949	Invalid
	*X*1.5	1.337	0.416	Invalid
*X*2 (movement characteristics)	*X*2.1	1.180	0.537	Invalid
	*X*2.2	1.180	0.207	Invalid
Y (chance of accident)	*Y*1	1.622	0.078	Invalid
	*Y*2	1.461	0.040	Valid
	*Y*3	1.126	0.622	Invalid
	*Y*4	1.032	0.283	Invalid
	*Y*5	1.285	0.033 (0.055)*	Invalid
	*Y*6	1.052	0.626	Invalid
	*Y*7	1.162	0.789	Invalid

*Note: * – after the second modification became insignificant.*

Variable *X*1 (socio economic characteristics) with all indicators have VIF smaller than five and valid items with a p-value reduced than 0.05. Then, there is one valid item out of 5 in variable *X*1. (Then the items *X*1.1, *X*1.2, *X*1.4, and *X*1.5 are eliminated).

The *X*2 variable (Movement Characteristics) with all indicators has a VIF smaller than five and a valid item with a p-value smaller than 0.05. Therefore, there are no valid items out of 2 in variable *X*2. (Therefore, variable *X*2 is eliminated from this model).

Reflective indicators (variable X3).

The tests to measure the Outer model with insightful indicators will include assessments for convergent validity, discriminant validity, composite reliability, average variance extracted (AVE), and Cronbach's Alpha Table 5 shows the results of the outer model results from reflective indicators (after modification).

Table 5

Outer model measurement results from reflective indicators (after modification)

Latent variable	Item	Outer loading	Description	Cronbach's Alpha	Composite reliability	AVE
$X3$ (driver behavior characteristics)	$X3.1$	0.547	Invalid	0.841	0.869	0.632
	$X3.2$	0.841	Valid			
	$X3.3$	0.933	Valid			
	$X3.4$	0.806	Valid			

From Table 5, it is evident that not all indicators were initially deemed valid; indicators are considered valid if their outer loading value exceeds 0.700. Some items are most dominant in reflecting these variables in each variable. The results are as follows.

After eliminating $X3.1$, the greatest main indicator in variable $X3$ (driver behavior characteristics) is $X3.3$, namely Ordinary Violation behavior with the highest loading factor of 0.933.

Referring to the Table 5, it is evident that all indicators exhibit construct-forming cross-loading values greater than those on other variables. Additionally, the loading factor value of the $X3$ variable indicators remains higher than that of the $X3$ variable itself.

After only valid items, the Cronbach's Alpha value of variable $X3$ of 0.844 is more significant than 0.7, indicating that variable $X3$ (Characteristics of driver behavior) is reliable. The variable $X3$ (driver behavior characteristics) exhibits a composite reliability figure of 0.902, surpassing the threshold of 0.7, indicating high reliability. Furthermore, the discriminant validity, as assessed by the Average Variance Extracted (AVE) value, confirms the validity of variable $X3$, with a value of 0.755, exceeding the acceptable threshold of 0.5.

5. 1. 3. Goodness of fit for two-wheeled vehicles

Table 6 describes the coefficient of determination (R-square) obtained from model 1 which assesses the effect of variables $X1$ (socio economic characteristics) and $X3$ (driver behavior characteristics) on Y (accident probability).

Table 6

Table of determination coefficient results

Influence	R Square
$X1$ (socio-economic characteristics, $X3$ (driver behavior characteristics))$\rightarrow Y$ (Chance of Accident)	0.102

The coefficient of determination (R-square) obtained from model 1 of variables X1 (socio economic characteristics) and X3 (driver behavior characteristics) on Y (Accident Opportunities), is 0.102. This indicates that 10.2 % of the variance in variable Y (Accident Opportunities) can be explained by the independent variables X1 and X3, while the remaining 89.8 % of the variance is influenced by other variables not examined in this study.

Effect Size (F^2).

Table 7 indicated the F-square value represents the effect size or proportion of the variance in endogenous variables explained by exogenous variables. F-square coefficients fall into three categories: small (0.02 to 0.15), medium (0.15 to 0.35), and large (greater than 0.35).

Table 7

Effect size result table

Exogenous	Model 1 (Y)		Goodness of fit index (GoF)
	F square	Effect	
$X1$ (socio-economic characteristics)	0.054	Small	0.319
$X3$ (driver behavior characteristics)	0.023	Small	

It was found that the F-square value in this study had a small effect size for socio economic characteristics obtained a value of 0.054 and for driver behavior characteristics obtained a value of 0.023.

The goodness of Fit model testing is carried out to see the model's overall accuracy by multiplying the average coefficient of determination by the average communality value (AVE). The calculation result for the Goodness of Fit Index (GoF) is 0.319, indicating that the model's accuracy falls within the medium category (0.25–0.35).

Hypothesis test.

The constants indicating of Table 8 is the influence of one covert variable on additional. An effect is deemed substantial uncertainty the p-value is less than 0.05, while it is considered irrelevant if the p-value exceeds 0.05. The computational outcomes, facilitated by SmartPLS software, yielded the following results:

Table 8

Effect results with T-statistics

Influence	Path coefficient	T statistics	p-values	Description
$X1 \rightarrow Y$	–0.234	3.353	0.001	Significant
$X3 \rightarrow Y$	0.153	2.276	0.023	Significant

Variable $X1$ (socio economic characteristics) exerts a negative and significant influence on variable Y (accident opportunities), as evidenced by T-statistics values exceeding the serious value ($3.353>1.96$) and p-values less important than α ($0.001<0.050$). The negative coefficient suggests that an increase in variable $X1$ (socioeconomic characteristics) significantly reduces variable Y (accident probability).

On the other hand, variable $X3$ (driver behavior characteristics) demonstrates a positive and significant impact on variable Y (accident chance), with T-statistics values surpassing the serious value ($2.276>1.96$) and p-values lesser than α ($0.023<0.050$). The positive coefficient indicates that an increase in variable $X3$ (driver behavior characteristics) significantly escalates variable Y (accident chance).

5. 1. 4. Goodness of fit for four-wheeled vehicles

Table 9 describes the coefficient of determination (R-square) obtained from model 1 which assesses the effect of variables $X1$ (socio economic characteristics) and $X3$ (driver behavior characteristics) on Y (Accident Probability).

The coefficient of determination (R-square) gained from model 1, namely the effect of variables $X1$ (socioeconomic characteristics) and $X3$ (driver behavior characteristics) on Y (accident opportunities) of 0.217, so that variable Y (accident opportunities) can be explained by independent variables $X1$ and $X3$ by 21.7 % and other variables outside this study influence the remaining 78.3 %.

Table 9

Table of determination coefficient results

Influence	R Square
$X1$ (socioeconomic characteristics), $X3$ (driver behavior characteristics)$\rightarrow Y$ (Chance of Accident)	0.217

Effect size (F^2).

The F-square value represents the effect size or proportion of the variance in endogenous variables explained by exogenous variables. F-square coefficients fall into three categories: small (0.02 to 0.15), medium (0.15 to 0.35), and large (greater than 0.35) which is shown in Table 10.

Table 10

Effect size&goodness of fit index (GoF) result table

Exogenous	Model 1 (Y)		Goodness of fit index (GoF)
	F-square	Effect	
$X1$ (socio economic characteristics)	0.149	Small	0.466
$X3$ (driver behavior characteristics)	0.102	Small	

It was found that the F-square value in this study had a small effect size for socio economic characteristics obtained a value of 0.149 and for driver behavior characteristics obtained a value of 0.102.

The goodness of Fit model challenging is carried out to realize the model's complete accuracy by multiplying the average coefficient of determination by the average communality value (AVE). The GoF calculation result is 0.466, so it can be concluded that the model's accuracy is in the medium type (0.25–0.35).

Hypothesis testing.

This section evaluates the coefficients or parameters indicating the influence of one latent variable on another. An effect is considered significant if the p-value is less than 0.05, while it is considered insignificant if the p-value exceeds 0.05. The computational outcomes, facilitated by SmartPLS software, yielded the following results which is shown in Table 11.

Table 11

Table of effect results with T-statistics

Influence	Path coefficient	T-statistics	p-values	Description
$X1 \rightarrow Y$	–0.343	2.841	0.005	Significant
$X3 \rightarrow Y$	0.284	3.218	0.001	Significant

Variable $X1$ (socio economic characteristics) exhibits a negative and significant influence on variable Y (accident opportunities), with T-statistics values exceeding the serious value (2.841>1.96) and p-values lesser than α (0.005<0.050). The negative coefficient suggests that an increase in variable $X1$ (Socioeconomic characteristics) can significantly decrease variable Y (accident probability).

Variable $X3$ (driver behavior characteristics) demonstrates a positive and significant influence on variable Y (accident chance), with T-statistics values surpassing the dangerous value (3.218>1.96) and p-values minor than α (0.001<0.050). The positive coefficient implies that an increase in variable $X3$ (driver behavior characteristics) can notably elevate variable Y (accident chance).

5. 2. Traffic accident probability model

5. 2. 1. Two-wheeled vehicle

The path diagram image shows the relationship between the path coefficients in the structural model and the weighted values of the manifest variables in the measurement model when considering the external factors of the study in Fig. 2 below.

From the diagram provided above, the model equation can be derived as follows:

$$Y = -0.234\ X1 + 0.153\ X3 + ei2;\ R^2 = 0.102. \qquad (1)$$

A structural equation model (SEM) that illustrates the relationships between different latent variables and observed indicators. Here's a breakdown of the model.

The latent variable $X1$, representing socio-economic characteristics, is determined through two observed indicators, $X1.1$ and $X1.2$, showing significant factor loadings that indicate how strongly each indicator represents the latent variable.

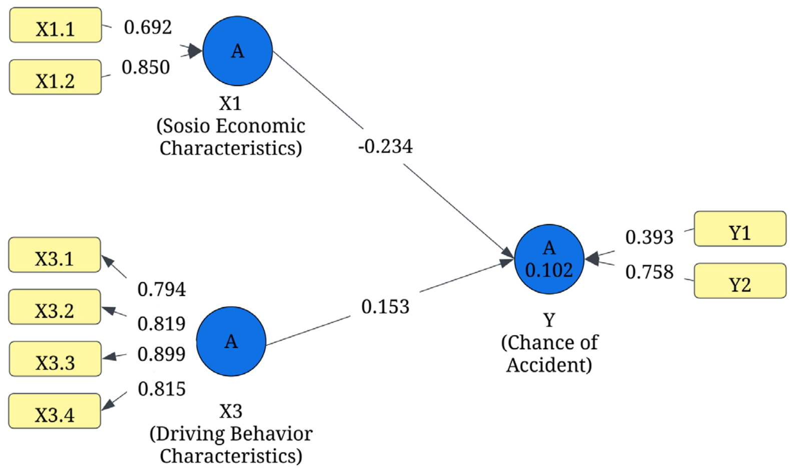

Fig. 2. Structural model diagram after model modification

The latent variable $X2$, indicative of driving behavior characteristics, is quantified by four observed indicators: $X3.1$, $X3.2$, $X3.3$, and $X3.4$. Each has a high factor loading, demonstrating a strong correlation to the latent variable.

The latent variable Y, reflecting the chance of an accident, is directly influenced by the variables $X1$ and $X2$ and subsequently influences the outcomes observed through $Y1$ and $Y2$, with their respective loadings illustrating the degree of influence Y has on these outcomes.

Path Coefficients: The path from $X1$ to Y (chance of accident) has a coefficient of –0.234, indicating a negative influence of socio-economic characteristics on the chance of an accident; The path from $X2$ to Y has a coefficient of 0.153, indicating a positive influence of driving behavior characteristics on the chance of an accident.

Variance Explained: The latent variable Y (chance of accident) has a variance explained value of 0.102, meaning that 10.2 % of the variance in the chance of an accident is explained by the socio-economic and driving behavior characteristics.

This model helps to understand how socio-economic and driving behavior characteristics impact the chance of an accident and how this, in turn, affects certain outcomes (*Y*1 and *Y*2). The factor loadings and path coefficients provide insights into the strength and direction of these relationships.

5. 2. 2. Four-wheeled vehicle

The path diagram image shows the relationship between the path coefficients in the structural model and the weighted values of the manifest variables in the measurement model when considering the external factors of the study in Fig. 3 below.

Based on the picture above, the model equation is obtained as follows:

$$Y = -0.343X1 + 0.284X3 + ei2;\ R^2 = 0.102. \quad (2)$$

A structural equation model (SEM) that illustrates the relationships between different latent variables and observed indicators. Here's a breakdown of the model.

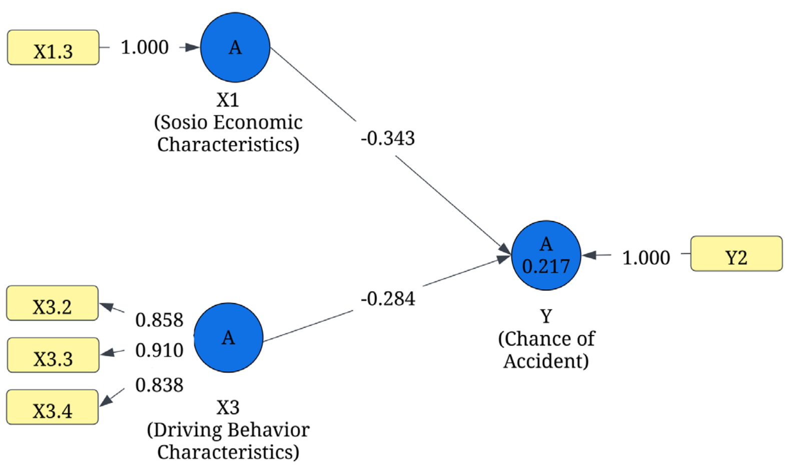

Fig. 3. Structural model diagram after model modification

The latent variable *X*1, representing socio-economic characteristics, is accurately determined by a single observed indicator, *X*1.3, which shows a perfect factor loading of 1.000, demonstrating an exact measurement.

The latent variable *X*2, indicative of driving behavior characteristics, is quantified through three observed indicators (*X*3.2, *X*3.3, *X*3.4). Each indicator has a strong factor loading, signifying robust measurements for capturing aspects of driving behavior.

The latent variable *Y*, reflecting the chance of an accident, is influenced by both *X*1 and *X*2. The outcome variable *Y*2 is directly measured with a factor loading of 1.000, confirming it as a precise measure of accident probability.

Path Coefficients. The path from *X*1 to *Y* has a coefficient of –0.343, suggesting a negative influence of socio-economic characteristics on the chance of an accident. The path from *X*2 to *Y* has a coefficient of –0.284, also indicating a negative influence of driving behavior characteristics on the chance of an accident.

Variance Explained. The latent variable *Y* (Chance of Accident) has a variance explained value of 0.217, meaning that 21.7 % of the variance in the chance of an accident is explained by the socio-economic and driving behavior characteristics.

This model provides insight into how socio-economic factors and driving behaviors collectively impact the probability of accidents. The negative path coefficients imply that better socio-economic and driving behavior characteristics are associated with a reduced chance of accidents.

6. Discussion of vehicle accident probability model

The analysis conducted using SEM (structural equation modeling) shows that the main factors contributing to accidents are socio-economic variables (*X*1) found in Tables 2, 7 and driving behavior (*X*3) found in Tables 3, 8. Which has consistently identified human factors, specifically driving behavior, as the principal reason of accidents. In this study, the predominant pointer within the driving behavior variable (*X*3) is ordinary violations (*X*3.3), with a weight of 0.903 for two-wheel vehicles and 0.933 for four-wheel vehicles. Therefore, the higher the frequency of driver infractions such as disregarding speed limits, running red lights, improper overtaking, driving under the influence, and failure to maintain a safe distance from other vehicles, the greater the likelihood of accidents.

This research employs the SEM (structural equation modeling) method using SmartPLS software. The advantages of SEM contain its capability to holder intricate dealings between variables. Variables container theoretical or unobservable. The situation approximations all coefficients popular model simultaneously, enabling the calculation of the significance and strength of specific relationships within a comprehensive model. Moreover, SEM accounts for multicollinearity and eliminates measurement errors, ensuring the validity of coefficients [21]. Previous research [22] that utilized Lisrel to analyze driving behavior yielded similar variables regarding the causes of accidents.

However, the study immobile relies on the second-generation SEM (structural equation modeling) technique, with SmartPLS package chosen for its user-friendly interface and fewer assumptions required. One limitation of this study is the absence of analysis of numerous variables related to the causes of accidents. The author encourages upcoming research to incorporate a broader range of variables associated with accident causation to enhance validity and address existing limitations. Additionally, advancements in analysis methods, such as utilizing third-generation SEM, could further improve the robustness of the research findings. Furthermore, exploring similar characteristics in other fields could provide valuable insights and broaden the scope of research in this area.

7. Conclusions

1. Characteristics of motor vehicle drivers on roads in Jayapura City based on analysis using SEM-PLS for two-wheeled vehicles on socio-economic variables (*X*1), which have dominant indicators, namely age (*X*1.1) through the uppermost factor weight of 1,793. The driver behavior variable (*X*3)

has the dominant indicator, namely ordinary violation ($X3.3$), with the highest weight of 0.903. For four-wheeled vehicles on socio-economic variables ($X1$), which have the dominant indicator, namely the last education ($X1.3$) by the uppermost factor weight of 1.246. The driver behavior variable ($X3$), which has the dominant indicator, namely ordinary violation ($X3.3$), has the highest weight of 0.933.

2. The model of accident opportunities involving motorized vehicles on road sections in Jayapura City for two-wheeled vehicles is $Y=-0.234X1+0.153X3+ei2$; $R^2=0.102$ where the highest path coefficient is found in the driver behavior variable ($X3$) of 0.153. And for four-wheeled vehicles, $Y=-0.343X1+0.284X3+ei2$; $R^2=0.217$ where the highest path coefficient is found in the driver behavior variable ($X3$) of 0.284. The research results show that the possibility of traffic accidents is caused by driver behavior. From this model, appropriate handling steps need to be taken, so that the problem of motorbike accidents can be resolved properly, such as adding billboards as a warning to reduce vehicle speed, providing shock markers so that drivers remain focused on driving, and providing warnings to take action. rest if they travel long distances.

Conflict of interest

The authors declare that they have no conflicts of interest regarding this research, whether financial, personal, authorship-related, or otherwise, that could influence the research and the results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Acknowledgments

I am deeply thankful to Dr. Ir. M. Zainul Arifin, MT, and Dr. Fauzul Rizal Sutikno for their invaluable guidance and insightful advice throughout my research. Additionally, I encompass my gratitude to my parents, siblings, and friends meant for their unwavering care, encouragement, and prayers through the writing process of this article.

References

1. Van Elslande, P., Elvik, R. (2012). Powered two-wheelers within the traffic system. Accident Analysis & Prevention, 49, 1–4. https://doi.org/10.1016/j.aap.2012.09.007
2. Retallack, A. E., Ostendorf, B. (2020). Relationship Between Traffic Volume and Accident Frequency at Intersections. International Journal of Environmental Research and Public Health, 17 (4), 1393. https://doi.org/10.3390/ijerph17041393
3. Pervez, A., Lee, J., Huang, H. (2021). Identifying Factors Contributing to the Motorcycle Crash Severity in Pakistan. Journal of Advanced Transportation, 2021, 1–10. https://doi.org/10.1155/2021/6636130
4. Suraji, A., Djakfar, L., Wicaksono, A., Marjono, M., Putranto, L. S., Susilo, S. H. (2021). Analysis of intercity bus public transport safety perception modeling using conjoint. Eastern-European Journal of Enterprise Technologies, 4 (3 (112)), 36–42. https://doi.org/10.15587/1729-4061.2021.239255
5. Najmy, A., Dewi, R. S., Ciptomulyono, U. (2018). Identifikasi Pengaruh Perilaku terhadap Tingkat Kecelakaan Lalu Lintas dengan Stuctural Equation Modeling (SEM). Internasional Riset & Teknologi Teknik (IJERT).
6. Bathan, A., de Ocampo, J., Ong, J., Gutierrez, A. A., Seva, R. R., Mariano, R. S. (2018). A predictive model of motorcycle accident involvement using structural equation modeling considering driver personality and riding behavior in Metro Manila. Proceedings of the International Conference on Industrial Engineering and Operations Management, 1783–1804. Available at: https://animorepository.dlsu.edu.ph/faculty_research/442
7. Putri, F., Arifin, M., Djakfar, L. (2022). Prediction model of motorcycle accident in economic and driving behaviour factors. Eastern-European Journal of Enterprise Technologies, 4 (3 (118)), 27–33. https://doi.org/10.15587/1729-4061.2022.263651
8. Hukom, F. S., Djakfar, L., Arifin, M. Z. (2023). Model Prediksi Kecelakaan Kendaraan Sepeda Motor pada Ruas Jalan di Kota Ambon. Rekayasa Sipil, 17 (2), 217–222. https://doi.org/10.21776/ub.rekayasasipil.2023.017.02.14
9. Miftahulkhair, M., Arifin, M. Z., Sutikno, F. R. (2024). Revealing the impact of losses on flexible pavement due to vehicle overloading. Engineering Technological Systems, 2 (1 (128)), 55–63. https://doi.org/10.15587/1729-4061.2024.299653
10. Chavan, E., Roopa, M. (2020). Automatic crash guard for motorcycles. International Journal of Electrical Engineering and Technology (IJEET), 11 (2), 17–26. Available at: https://sdbindex.com/Documents/index/00000003/00000-04007
11. Machsus, M., Sulistio, H., Wicaksono, A., Djakfar, L. (2014). Generalized linear and generalized additive models in studies of motorcycle accident prediction models for the north-south road corridor in Surabaya. The 17thFSTPT International Symposium, 976–986. Available at: https://jurnal.unej.ac.id/index.php/PFSTPT/article/view/2921/2347
12. Shah, S., Ahmad, N., Shen, Y., Pirdavani, A., Basheer, M., Brijs, T. (2018). Road Safety Risk Assessment: An Analysis of Transport Policy and Management for Low-, Middle-, and High-Income Asian Countries. Sustainability, 10 (2), 389. https://doi.org/10.3390/su10020389

13. Murphy, P., Morris, A. (2020). Quantifying accident risk and severity due to speed from the reaction point to the critical conflict in fatal motorcycle accidents. Accident Analysis & Prevention, 141, 105548. https://doi.org/10.1016/j.aap.2020.105548
14. Mohamed, M., Bromfield, N. F. (2017). Attitudes, driving behavior, and accident involvement among young male drivers in Saudi Arabia. Transportation Research Part F: Traffic Psychology and Behaviour, 47, 59–71. https://doi.org/10.1016/j.trf.2017.04.009
15. Zhang, G., Yau, K. K. W., Chen, G. (2013). Risk factors associated with traffic violations and accident severity in China. Accident Analysis & Prevention, 59, 18–25. https://doi.org/10.1016/j.aap.2013.05.004
16. Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. Ecological Processes, 5 (1). https://doi.org/10.1186/s13717-016-0063-3
17. Hassan, H. M. (2015). Investigation of the self-reported aberrant driving behavior of young male Saudi drivers: A survey-based study. Journal of Transportation Safety & Security, 8 (2), 113–128. https://doi.org/10.1080/19439962.2015.1017782
18. Yamin, S., Kurniawan, H. (2009). SPSS Complete: teknik analisis statistik terlengkap dengan software SPSS. Jakarta: Salemba Infotek, 328.
19. Sholihin, M., Ratmono, D. (2021). Analisis SEM-PLS dengan WarpPLS 7.0: untuk hubungan nonlinier dalam penelitian sosial dan bisnis. Yogyakarta, 320.
20. Hair Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E. (2010). Multivariate Data Analysis. Pearson. Available at: https://www.drnishikantjha.com/papersCollection/Multivariate%20Data%20Analysis.pdf
21. Jung, S., Xiao, Q., Yoon, Y. (2013). Evaluation of motorcycle safety strategies using the severity of injuries. Accident Analysis & Prevention, 59, 357–364. https://doi.org/10.1016/j.aap.2013.06.030
22. Shaheed, M. S., Gkritza, K. (2014). A latent class analysis of single-vehicle motorcycle crash severity outcomes. Analytic Methods in Accident Research, 2, 30–38. https://doi.org/10.1016/j.amar.2014.03.002