Recent Trends in Civil Engineering and Built Environment Vol. 4 No. 2 (2023) 172-179

RTCEBE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/rtcebe
e-ISSN :2773-5184

Investigation on the Correlation Between the Road Accidents and Driver Behaviours Through Observations and Statistical Analysis Along Batu Pahat – Kluang

Muhammad Syahrunnizam Jamaludin[1], Nasradeen. A. Khalifa[1*]

[1]Faculty of Civil Engineering and Built Environment,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

*Senior Lecturer, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia

DOI: https://doi.org/10.30880/rtcebe.2023.04.02.020
Received 06 January 2022; Accepted 15 January 2023; Available online 20 July 2023

Abstract: Batu Pahat – Kluang route has the highest occurrence of accidents in Malaysia. According to Jabatan Siasatan dan Trafik Bukit Aman, the ADT (Annual Daily Traffic) in 2019 for cars is 36,704 and the most accidents occurred during that year with 573,341. The objective of this study is to investigate the correlation between behaviors and road accidents as well as understanding the patterns of the behavior. By using questionnaires, the data then analuzed using parametric methods via SPSS. The highest behavior act by the drivers that were using FT050 route is behavior that related to driving skills and the lowest is the behavior that related to interactions among drivers on the road. Therefore, further research needs to be implemented in order to gain more complex data using the helps of Artificial Intelligence (AI) so that the behavior can be analyzed more accurately and provide future solutions.

Keywords: FT050, SPSS, Parametric, Driving Skills, behavior, interactions, patterns

1. Introduction

In modern worlds, transportation plays an important role in our daily life. Despite the convenience that it served humanity; problems arise when there are too many vehicles on the road at the same time. The most critical problem is road accidents. The World Health Organization stated that the number of deaths caused by road accidents or road accidents in 2016 is 1,350,000 people [6]. In Malaysia, as reported by the Ministry of Transport Malaysia, the total fatalities in 2016 have reached a peak of 7152 death cases with the lowest in 2019 being 6,167 death cases. But the number of cases related to road accidents continues to increase according to statistical data from 2010 until 2019. Based on a newspaper article dated 28th May 2019 by New Straits Time [7], Batu Pahat – Kluang route has the highest occurrence

*Corresponding author: nasradeen@uthm.edu.my

of accidents in Malaysia. According to Jabatan Siasatan Dan Trafik Bukit Aman [8], the ADT (Annual Daily Traffic) in 2019 for cars is 36,704 and the most accidents occurred during that year with 573,341.

There are many causes that can lead to road accidents, such as improper driving, out-of-date services, technical problems, etc. According to Underwood [9], states that, the highest factor led to road accidents are human factors. The most common method used is psychoanalytical methods. It is considered the most suitable method to analyze the positive and negative elements of the behaviors that reflected by the conscious mind of a person's conscious [10]. Past studies show that attitudes towards traffic safety have been shown to correspond to belligerent driving behavior, fast driving, and self-reported accident collusion [1]

DBQ (Driver Behaviour Questionnaire) is a questionnaire that has been developed as a method to measure driving behaviours. The main advantage of survey analysis is that in a single questionnaire an immense amount of data can be collected that can later be analysed through statistical analysis [2]. According to Montuori [11] he stated that most of the road accidents that occur in Italy were due to the use of mobile phones while driving. Another study that has been conducted in China also stating that most of the road accidents that occurred were believed to be responsible for autonomous drivers with 92% of the accidents. In that study, the main drivers' behaviors that contribute to road accidents resulting in fatalities were speeding, reckless driving, wrong lane driving, unlicensed driver, and driving under the influence of alcohol [3]. In Vietnam, mobile phones contribute to 20.3% reported crashes [4]. In addition to that the, in Japan also found that influence of alcohol, driving at inappropriate speeds, and using mobile phones while driving also contribute to road accidents [5]. All of the past research that has been done showing the contribution of human driving behaviors can cause unprecedented road accidents. It can be either disruptions towards humans functional by consuming hazardous chemicals or sleepiness, or bad driving behaviors that have been normalized among road users.

2. Assessment of Drivers Behaviors

Parametric testing will be conducted by using SPSS (Statistical Package for Social Sciences). Therefore, questionnaires will be distributed throughout various social media platform via google form. This helps in terms of contactless method due to Covid-19 restrictions. Basically, the data collection must achieve the targeted quota for a strong analysis which is 379 samples. Other than that, the data that have been collected must go through data cleaning process and the determination of the normality test. This is to ensure that the data has no missing numbers and exclude the non-reliable data samples. The normality test to determine the type of data analysis that will be used such as parametric or non-parametric test. Next, modified HIRARC analysis model will be used to analysis the ranks of behaviors from the most occurrence to the lowest.

3. Samples and Methods

The samples and methods will be discussed in this section to understands the process of obtaining the data and the analyzation method.

3.1 Samples

Specifics settings and criteria has been made to make the data collection is more focus and more systematic. The criteria are as follows:

- Drivers who use/used FT050 route (Batu Pahat – Ayer Hitam – Kluang)
- Vehicles with 2-axles or 3 and 4 wheels only

3.2 Methods

The methods that will be used in this study focused on data analysis and modified HIRARC model. The steps are as follows:

a. Pilot study

In pilot study, there are several steps that needs to be done. The cleaning of data is mandatory to exclude the non-reliable data such as missing numbers and outliers of the samples. Upon completion, the data then will be going through the normality test to determine whether the data is normal or not normal. If the data is normal, then it can be analyzed through parametric analyzation methods.

b. Descriptive analysis

Descriptive analysis is the analysis towards the frequency. Demographic data will be analyzed to recognize the pattern of the samples. In addition to that, the computed variables will also be analyzed to produce the normal distribution chart and the histogram line to understand the pattern of the behaviors for further analysis. The results are shown in forms of graph, pie chart, and tables.

c. Correlation

In correlation analysis, the computed variable was tested by using Pearson's correlation to identify if there are relationships between the computed variables. This helps to understand whether that particular behavior has a significance relationship towards another computed variable.

d. HIRARC modified

To achieve the last objective of this project, statistical analysis is required to be provided for determine which types of behaviors contribute the most to road accidents and rank them accordingly. The analysis that will be used is the modified HIRARC table.

3.3 Equations

The equations that will be used in this project is the sample size calculation to determine the number of samples that needed for achieving a strong and reliable data to analyze. The method that will be used is Normal Distribution method.

$$x = Z\left(\frac{c}{100}\right)^2 r(100 - r) \quad (1)$$

$$n = N\,x/((N-1)E^2 + x) \quad (2)$$

Where n is sample size, E is the margin of error (usually 0.05), N is the population size, r is the responses fraction (usually 0.5) and Z(c/100) is the critical value for the confidence level c.

Cronbach's alpha is a measure of the internal consistency or reliability of a set of items. A high value of alpha considered a higher reliability of the items.

$$\alpha = \frac{N\bar{c}}{\bar{v}+(N-1)\bar{c}} \quad (3)$$

Where N is the number of samples, $\bar{c}$ is the average inter-item covariance among the items and $\bar{v}$ is the average variance.

4. Results and Discussion

The results of the study would be discussed as mentioned in 3.2 in order to understand fully about the findings.

4.1 Results

a. Pilot study

Before any data can be processed, the reliability of the data must be done via Cronbach' Alpha (Eq.1). By using KNIME Analytical Platform, the value obtained is 0.94 which considered as excellent. Then, the data cleaning process can be done, and the results are shown as in Figure 1. The samples that are

located out of the box plot range will be considered as unreliable and excluded from the process of the data.

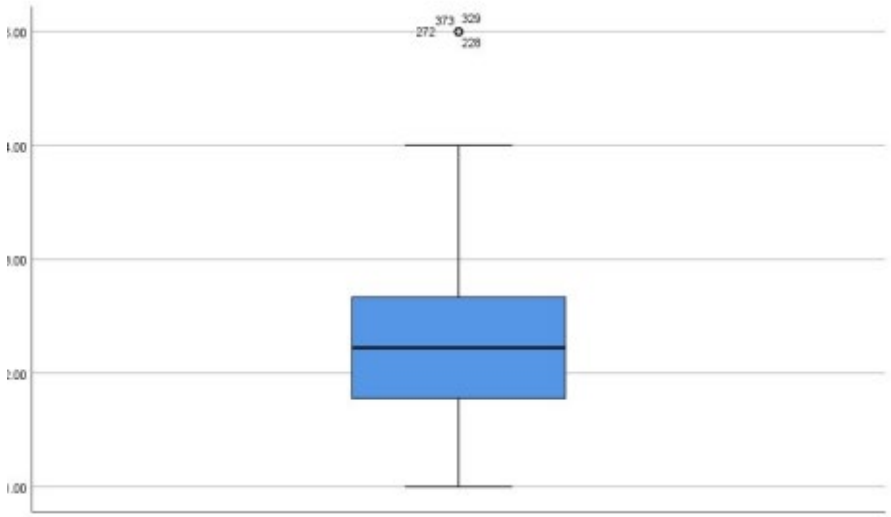

Figure 1: Outliers for computed variable driving skill behaviour

Next, the data will be put through the normality test to determine whether the data is normal or not. Table 1 shows the summary of the normality test. As the results proven the data is normal, then parametric analysis will be the method that is suitable for the analyzation of the data.

Table 1: Summary of normality test

Compute Variable	Kolmogorov-Smirnov Sig. (P>0.05)	Skewness (-1.96<P<1.96)	Remarks
Section B	0.00	0.99	Normal
Section C	0.00	1.052	Normal
Section D	0.00	0.22	Normal
Section E	0.00	1.08	Normal

From Table 1, the Kolmogorov-Smirnov Sigma value for all the computed variables is 0, therefore the skewness will substitute the condition of the normality of the data. From sections A, B, C, and D, all the skewness value are within the range of -1.96 and +1.96. therefore, all the data will be considered normal thus, parametric analysis will be used for the data evaluation.

b. Descriptive analysis

Descriptive analysis gives understandings towards the patterns of the data. Therefore, the important result for this study is to analyze which behavior has exceed the mean value of the computed variable. Table 2 gives the summary towards the results that have been obtained.

Table 2: Summarization of the comparison between compute variable and item/behaviour.

Computed Variable of Behaviors	Mean Value	Item/Behavior > Mean Value
Driving Skill	2.34	B2, B3, B4, B6
Traffic Law Comprehension	2.33	C1, C2, C3, C6
Driver's Consciousness	2.65	D1, D2, D3
Driver's Interaction With Each Other	2.12	E1, E2

From Table 2, the mean value of Driving Skill is 2.34 which has 4 behavior that exceed it which are B2, B3, B4 and B6. As for the Traffic Law Comprehension, the mean value is 2.33 which has 4 behavior that exceed it which are C1, C2, C3 and C4. For Driver's Consciousness, the mean value is 2.65 which has 3 behaviors that exceeds it which are D1, D2 and D3. Last one is the Driver's Interaction With Each Other which has the mean value of 2.12 have 2 behaviors that exceeds it which are E1 and E2.

c. Correlation

In correlation analysis, the computed variable was tested by using Pearson's correlation to identify if there are relationships between the computed variables. This helps to understand whether that particular has a significance relationship towards another computed variable.

Table 3: Correlation table between computed variables

Correlations					
		Driving Skills	**Law Comprehension**	**Drivers Consciousness**	**Drivers Interactions With Each Other**
Driving Skills	**Pearson Correlation**	1	.605**	.683**	.578**
	Sig. (2-tailed)		0.000	0.000	0.000
	N	386	386	386	386
Law Comprehension	**Pearson Correlation**	.605**	1	.656**	.761**
	Sig. (2-tailed)	0.000		0.000	0.000
	N	386	386	386	386
Drivers Consciousness	**Pearson Correlation**	.683**	.656**	1	.620**
	Sig. (2-tailed)	0.000	0.000		0.000
	N	386	386	386	386
Drivers Interactions With Each Other	**Pearson Correlation**	.578**	.761**	.620**	1
	Sig. (2-tailed)	0.000	0.000	0.000	
	N	386	386	386	386

From Table 3, there are significance relationship between the computed variables as the value of Sigma (2-tailed) is 0.00. To determine whether the relation is stronger, weaker, or no relations, the Pearson correlation value will be observed. The summary of the relations is as follows:

Table 4: Pearson correlation summary

Pearson Correlation				
Computed Variables	**Driving Skills**	**Traffic Law Comprehension**	**Drivers Consciousness**	**Drivers Interaction With Each Other**
Driving Skills		0.60	0.68	0.57

Traffic Law Comprehension	0.60		0.65	0.76
Drivers Consciousness	0.68	0.65		0.62
Drivers Interaction With Each Other	0.57	0.72	0.62	

Based on Table 4, as the Pearson's correlation value for all the relationships is between 0.50 and 1.00, it shows that all the correlations have strong relationships with each other.

d. HIRARC modified

To achieve the last objective of this project, statistical analysis is required to be provided for determine which types of behaviors contribute the most to road accidents and rank them accordingly. The analysis that will be used is the modified HIRARC table.

Table 5: Modified HIRARC model to determine behaviour ranks

Behaviour Identification			**Risk Analysis**			Rank
Behaviour	**Hazard**	**Effect**	**Likelihood**	**Severity**	**Risk**	
Driving Skills	Bad driving skills can lead to accidents because they interfere with others. For example, misjudge the incoming speed of other vehicles when turning can increase the chance of collision	Road accidents	2.34	3	7.02	1
Traffic Law Comprehension	Ignoring traffic laws can seriously lead to road accidents. For example, ignoring the red lights in traffic can lead to collisions with other transportation that have the right of way	Road accidents	2.33	3	6.99	2
Drivers Consciousness	Driving without paying full attention to the surroundings presents an instant danger that can lead to rad accidents. For example, using the phone while driving takes half of the attention to the phone usage and limited the reaction	Road accidents	2.65	2	5.3	3

	towards any road hazards.					
Drivers Interaction With Each Other	Showing ill intention while driving towards other can propose threat or dangers to road users. For example, showing the intention to harm a certain individual driver on the road might cause harm towards unrelated road users.	Road accidents	2.12	1	2.12	4

Based on Table 4.14, the analysis of ranks behavior has been done by using modified HIRARC model. The number 1 rank is for the driving skills behaviors that possess the highest value of risk, which is 7.02. This shows that the possibility of road accidents occurrence is the highest as majority of the drivers have problems regarding their driving skills behaviors. Followed by Traffic Law Comprehension behavior, which has the value of risk 6.99. This shows that there is still the possibility of road accidents even though the drivers show a positive attitude regarding the traffic law. At rank number 3, Driers Consciousness while driving has the risk value of 5.3 followed by rank number 4 which is behavior regarding Drivers Interaction With Each Other with the lowest value of risk, which is 2.12.

5. Conclusion

The types of behaviors shown by the drivers along the FT050 are divided into 4 main parts. The first part is behavior with respect to the drivers driving skills. Based on the analysis, the sample driving skills prove to be in positive or good behavior as the shape of the histogram chart has a more tendency at the left side of the graph with the mean value of 2.34. The next part of behavior is towards the traffic law comprehension of the drivers. The analysis shows that the drivers who use the route FT050 has positive behavior with respect to the traffic law. This is proven by the histogram chart, which has a tendency at a left side with the mean value of 2.33. The third part of the behavior concerns the consciousness of the drivers while driving. The results show that the behavior is positive as the histogram chart has a higher tendency at a left side with the mean value of 2.65. The last part of the behavior is about the interactions between other drivers that use the FT050 route. The analysis also shows that the behavior behaves in a positive way. Through analysis, the histogram chart has a more tendency at left side and with a mean value of 2.12.

Each part of the computed variable, there are some items that exceed the mean value of the computed variable. This means that the particular behavior has more occurrence in the drivers and might have more tendency to the road accidence contributions.

From Table 5, the statistical analysis that has been done via the model has produced the rankings of each behavior. Point measures are the value of the risk value. Higher risk value means higher occurrence of behavior.

Acknowledgement

The authors extend their thanks to Faculty of Civil Engineering and Built Environment (FKAAB), for their support during my research work.

References

[1] Pa'l Ulleberg, "Personality, attitudes and risk perception as predictors of risk driving behaviour among young drivers," *Scandinavian Journal of Psychology,* no. 43, pp. 1-2, 2002.

[2] Sybille Gumucio, "The KAP Survey Model," January 2011. [Online]. Available: https://www.spring-nutrition.org/publications/tool-summaries/kap-survey-model-knowledge-attitudes-and-practices. [Accessed 24 July 2021].

[3] Zhang. X., "Basic characteristics of road traffic deaths in China," *Public Health,* vol. I, no. 42, pp. 7-15, 2013.

[4] Truong. L. N., "Mobile phone use among motorcyclists and electric bike riders: a case study of Hanoi, Vietnam.," *Accident Analysis Preview,* 2016.

[5] International Transport Forum, "Transport Connectivity for Regional Intergration," OCED, Leipzig, 2019.

[6] World Health Organization, "Monitoring Health for The SDGs, Sustainable Developement Goals," *World Health Statistics,* pp. 34-35, 2018.

[7] A. Abbas, *Batu Pahat - Kluang Highest Rate Accidents,* Putra: New Straits Times, 2019.

[8] *Jabatan Siasatan Dan Trafik Bukit Aman,* 2019.

[9] G. Underwood, Theory & Transport Psychology, Theory and Appllication, Nottingham: Elsevier, 2005.

[10] Alberto Bucchi, C. S. V. V., "Social and Behavioral Sciences," *Traffic Psychology and Driver Behavior,* pp. 974-980, 2012.

[11] Montuori, P., "The uses of mobile phone while driving," *Accident Analysis and Prevention: Behavior and determinant analysis in one of the largest metropolitan area of Italy,* vol. 157, pp. 106-161, 2021.