2022 Rapid KL Bus Accident Analysis and Insights for Improved Safety and Operational Efficiency

*Fauzan Ghazi[1], Suraya Ya'acob[2], Mariam Mazlan[3], Nurlaili Balqis[4]

[1, 2]Faculty of Artificial Intelligence, [3]Azman Hashim International Business School, [4]Faculty of Social Sciences and Humanities, University Teknologi Malaysia

[1]ahmad.fauzan@graduate.utm.my, [2]suraya.yaacob@utm.my, [3]mariam90@graduate.utm.my, [4]nurlailibalqis@graduate.utm.my

Article history

Received:
19 Dec 2024

Received in revised form:
22 Dec 2024

Accepted:
24 Dec 2024

Published online:
27 Dec 2024

*Corresponding author
ahmad.fauzan@graduate.utm.my

Abstract

Rapid KL operates one of Malaysia's most extensive public transportation networks, facing challenges in ensuring passenger safety and minimizing operational disruptions. In 2022, multiple bus accidents resulted in injuries, fatalities, and significant financial losses. This study analyzes accident data from that year to identify high-risk routes, assess contributing factors like road conditions and driver behavior, and evaluate financial impacts. Using Tableau for visual analytics, this research uncovers critical patterns, providing actionable recommendations such as targeted driver training and route-specific interventions to enhance safety measures, optimize route management, and improve operational efficiency.

Keywords: *Public Transportation, Data Visualization, Visual Analytics, Rapid KL, Tableau*

1. Introduction

Public transportation is essential for urban mobility. It provides affordable and accessible travel options for millions of people, saving significant money spent on car loans, highway fees, and petrol. Rapid KL operates an extensive bus network in Malaysia across Kuala Lumpur, Selangor, Pahang, and Penang. However, the scale and complexity of managing such a large fleet bring significant challenges. This includes the safety of passengers, minimizing accidents, and accidents monetary losses. Bus accidents result in potential injuries and fatalities, impose substantial financial costs, disrupt services, and erode public confidence in the transportation system.

This research aims to dive deep into the accident cause of analysis for 2022 involving Rapid KL buses to identify the primary causes of bus accidents. It also aims to evaluate and assess risk and mitigation, primarily to evaluate the risk levels associated with different routes and bus groups. The final goal is to do a cost analysis and optimization to analyze the financial impact of bus accidents. To do so, the Tableau tool will be used to develop multiple data visualizations and

dashboards. It will help Rapid KL explore the cause of the loss and how they may optimize the route to minimize future accidents. Using Tableau with visual analytics may reveal hidden and unnoticed information, such as hazardous road locations, road conditions, bus driver status, route and bus group, etc. By providing good data visualization using Tableau, the Rapid KL management group can improve decision-making to minimize bus accidents that lead to financial loss and injuries.

Initial findings indicate that vehicle-related incidents, poor road conditions, and peak-hour operations contribute to accidents. These factors and route-specific risks lead to significant financial losses concentrated in specific periods. Insights derived from Tableau visualizations highlight the need for enhanced road infrastructure, targeted driver training, and optimized schedules during high-risk periods. The recommendations from this analysis provide a strategic pathway for improving Rapid KL's safety and operational outcomes, ensuring a safer and more reliable public transportation system.

2. Literature Review

Multiple studies emphasize the impact of driver behavior, roadway conditions, and environmental factors on bus accident severity. [1] identify that factors like driver age, history of traffic violations, and road profiles are critical in determining accident outcomes. Similarly, [2] highlights that poor road conditions and external factors like environmental conditions contribute to higher accident risks. Both studies suggest that integrating long-term accident data and using predictive models can help mitigate these risks by identifying hazardous routes and implementing preventive measures.

Several researchers focus on the role of vehicle design and operational strategies in enhancing safety. [3] emphasizes that optimizing bus design, such as improving seating arrangements and accommodating peak-hour capacities, can reduce passenger accidents, particularly during emergency braking. These findings align with recommendations by [4], who propose data-driven decision-making tools for optimizing public transport operations. Their research demonstrates how scenario analysis and visualization techniques can support route optimization and accident prevention, which is crucial for enhancing Rapid KL’s operational efficiency.

[5] and [6] discuss the importance of using visualization tools and techniques like Tableau to analyze accident data and uncover hidden patterns. Data-driven approaches and tools are essential to improving public transportation safety and efficiency. Each research shows that applying them can enhance decision-making processes by enabling the identification of accident-prone areas and optimizing routes. Overall, it is very relevant for Rapid KL’s strategy to identify high-risk routes and improve safety.

Understanding the financial implications of accidents is critical for Rapid KL. [7] investigates the economic burden of road accidents, highlighting how they can significantly impact GDP. Rapid KL's perspective underscores the need for a comprehensive cost-benefit analysis considering the financial losses associated with accidents. Incorporating economic factors into safety strategies allows for a more holistic approach to improving safety and operational efficiency.

Research on road safety measures provides valuable insights into reducing bus accident rates. [8] and [9] highlight effective interventions like road curve

improvement, black spot treatments, and targeted safety audits. These strategies have been proven to reduce accident frequencies and can be adapted for high-risk areas within the Rapid KL bus network. By focusing on preventive measures, these studies offer actionable insights for mitigating risks and enhancing overall route safety.

The literature converges on integrating data-driven analysis, targeted preventive strategies, and operational optimizations to improve Rapid KL safety. For Rapid KL, adopting these insights can lead to more informed decision-making, optimized route planning, and reduced accidents and associated financial losses. By leveraging data visualization tools like Tableau or Power BI and focusing on critical risk factors, Rapid KL can enhance its operational efficiency and passenger safety.

Table 1. Literature Review Table

Source	Summary	Key Element	Outcome
Risk Factors Affecting Fatal Bus Accident Severity: Their Impact On Different Types Of Bus Drivers [1]	Identifies vital risk factors like driver behavior, road conditions, and demographics affecting bus accident severity.	Accident Severity	Critical for assessing high-risk routes and implementing targeted interventions in Rapid KL.
The Impact Of Roadway Conditions Towards Accident Severity On Federal Roads In Malaysia [2]	Highlights how road conditions and external factors increase accident risks on Malaysian roads.	Road Conditions	Long-term data analysis can identify hazardous routes and guide safety measures in Rapid KL.
The Role Of Public Transport In Transport Safety And Public Safety [3]	Discusses safety improvements through bus design and operational strategies to reduce passenger accidents.	Safety Optimization	Enhanced vehicle design and capacity management can lower accident risks during peak hours.
Designing A Data Visualization And Analysis Tool For Supporting Decision-Making With Public Transportation Networks [4]	Proposes integrating data visualization and scenario analysis for better decision-making.	Data-Driven Decision Making	Visualization tools can optimize routes and improve safety strategies for Rapid KL.
Intra-City Traffic Data Visualization: A	Reviews visualization practices in traffic data	Traffic Data Visualization	Advanced visualization can improve traffic pattern analysis and

Systematic Literature Review [5]	analysis, identifying best practices and gaps.		decision-making in Rapid KL.
Road Crash Data Visualisation and Analytics Using Tableau for Mountainous Roadway Areas in Cameron Highlands, Malaysia [6]	Uses Tableau to analyze crash data and detect risk patterns.	Data Visualization	Tableau can be applied to Rapid KL routes to reveal hidden risks and enhance decision-making.
The Economic Burden of Road Traffic Accidents and Injuries: A Small Island Perspective [7]	Examines the economic impact of road traffic accidents on GDP.	Economic Impact	Understanding financial losses from accidents can guide Rapid KL in making cost-effective safety investments.
Road Accidents Analytics with Data Visualization: A Case Study in Shah Alam Malaysia [8]	Analyzes road accident data using visualization to pinpoint accident hotspots and patterns.	Accident Data Analysis	Visualization can help identify high-risk areas and refine safety strategies across Rapid KL routes.
Road Traffic Accident Data Analysis and Its Visualization [9]	Analyzes accident data to understand trends and improve safety management.	Traffic Safety	Data visualization techniques can help Rapid KL identify accident trends and refine safety measures.
Discussion on Application of Big Data Analysis in Improving Traffic Safety: Database of Traffic Accidents and Violation Reports [10]	Examines correlations between traffic violations and accidents using big data tools like Tableau.	Traffic Violations	Highlights data-driven strategies for traffic law enforcement and reducing accident rates.
Identification of Hotspot Areas For Traffic Accidents and Analyzing Drivers' Behaviors and Road Accidents [11]	Combines GIS and statistical models to locate traffic accident hotspots and analyze driver behaviors.	Acciden t Hotspots	Pinpoints high-risk areas for implementing effective traffic safety policies.

Using Computer Vision and Machine Learning to Identify Bus Safety Risk Factors [12]	Leverages video analytics and machine learning to analyze pedestrian and environmental factors in bus crashes.	Crash Factors	Suggests proactive measures to address high-risk areas and improve bus-pedestrian interactions.
Analysis of Risk Bus Driver Characteristics and Research on Risk Level Evaluation Methods for Bus Drivers [13]	Develops a risk level evaluation model using machine learning to assess bus driver risks.	Driver Risk Evaluation	Supports targeted management of high-risk drivers, improving overall bus safety.
Application of Intelligent Transportation System Data using Big Data Technologies [14]	Proposes a big data-driven decision support model for intelligent transportation systems.	Decision Support	Guides strategic transportation management and resolves infrastructure bottlenecks.

3. Methodology

The methodology employed in this analysis of Rapid KL bus accidents centers on the comprehensive examination of accident data from 2022. The Rapid KL team provided this data, ensuring access to first-hand and relevant information directly from the operational source. The study involved several crucial data handling and processing stages to ensure accuracy and reliability in the insights generated.

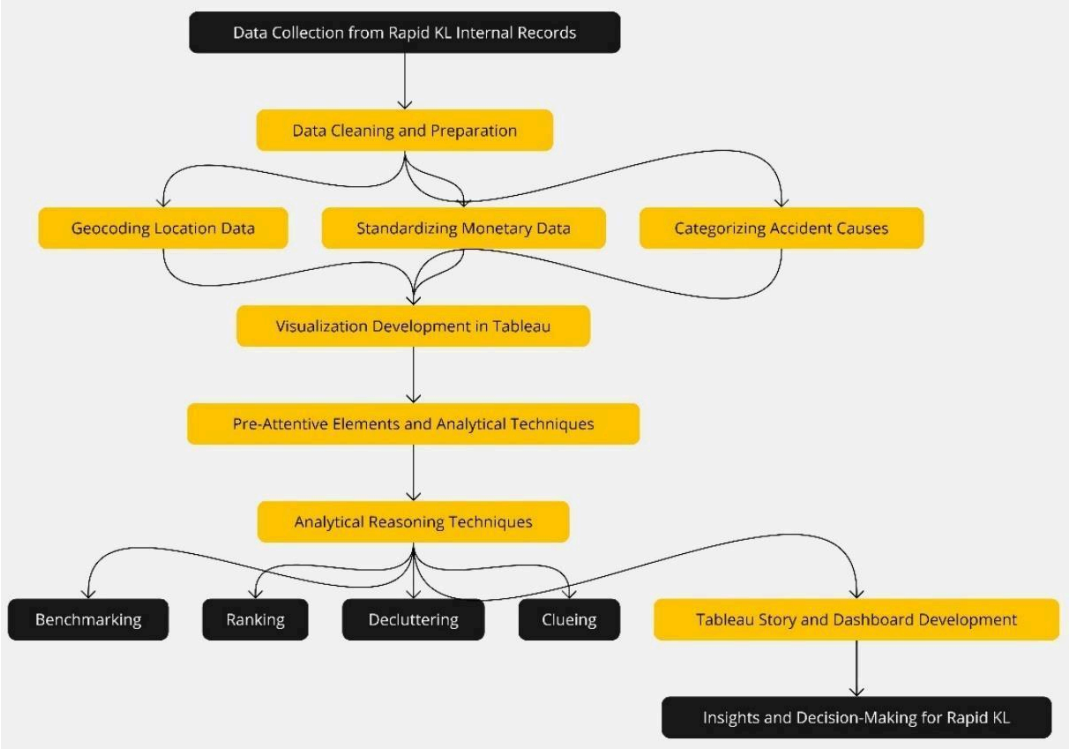

Figure 1. Research Design

3.1 Data Collection and Provision

The initial dataset was sourced from Rapid KL's internal records, which included detailed accounts of various accident-related parameters from 2022. This specificity in the dataset helped maintain focus on the most recent and applicable data for analyzing trends and patterns pertinent to current operational and safety standards.

3.2 Data Cleaning and Preparation

Given the complexity and volume of the data provided, significant efforts were put into cleaning and preparing the dataset for analysis. The data preparation process involved several vital steps. Geocoding Location Data Utilizing Google Apps Script with the Geocoding API, textual location data was transformed into precise geographical coordinates. This conversion was essential for enabling spatial analysis of the accidents, allowing for the identification of high-risk locations and trends over geographic distributions.

Extracting Numerical Values from Cost Data from the accident cost data, initially recorded as text, were processed to extract numerical values. This step was crucial for subsequent financial analyses and cost optimization studies, as it allowed for quantitative assessments and comparisons.

Categorizing Accident Causes with Python scripts were employed to parse and categorize textual descriptions of accident causes into definitive categories. This categorization helped streamline the analysis, making it possible to identify common themes and factors contributing to accidents.

These preparatory steps ensured the dataset was clean, well-structured, and optimized for in-depth analysis using data visualization tools. The cleaned and processed data served as a robust foundation for the subsequent phases of visualization and interpretation, facilitating a comprehensive analysis aimed at enhancing Rapid KL's operational safety and efficiency.

3.3 Data Visualization

The cleaned data was then visualized using Tableau to address business questions and deliver insights aligned with Rapid KL’s needs. The visualizations focused on pre-attentive elements and analytical reasoning techniques to enhance understanding and facilitate decision-making. Using pre-attentive elements, the visualizations strategically employed icons, colors, and spatial mapping.

Tableau was selected as the primary visualization tool for its robust integration capabilities, user-friendly interface, and advanced visualization options. Tableau proved effective for traffic accident data analysis and visualization of patterns and trends [6], [10]. Tableau was chosen over other tools like Power BI or Excel due to its ability to handle complex datasets and generate highly interactive visualizations. Tableau’s capability to merge data from diverse sources while providing intuitive drag-and-drop functionalities.

Various components of analytical reasoning, such as benchmarking, ranking, decluttering, and clueing, were applied.

3.4 Limitations and Biases

Recognizing the potential for biases such as underreporting minor accidents, this study incorporates cross-validation with external reports and uses data imputation

methods to account for missing entries. Limitations such as inconsistent recording of accident severity are mitigated through targeted cleaning and exploratory data analysis

4. Results and Discussion

The analysis of Rapid KL's 2022 bus accident data reveals critical patterns, factors contributing to accidents, and their associated financial impacts. The findings are discussed under three key areas: accident cause analysis, risk assessment and mitigation, and cost analysis and optimization. These are compared with findings from relevant studies to provide deeper insights and implications.

4.1 Accident Cause Analysis

The financial impact of bus accidents across 2022 amounted to RM2,010,397, with significant variations by month. The breakdown of costs revealed that accidents involving road conditions, particularly flyovers and steep inclines, accounted for RM825,291, underscoring the need for targeted road maintenance and clearer signage to prevent such incidents.

The predominant causes of accidents in 2022 were vehicle-related incidents (1,517 incidents, RM1,046,364.13 in losses), collisions with immovable objects, and poor road conditions. Interestingly, collisions with immovable objects, although less frequent, resulted in the highest financial impact of RM1,204,920.17. Environmental factors and technical failures contributed to less frequent but significant accidents. These findings align with studies [11], highlighting driver behavior, road conditions, and spatial accident hotspots as significant determinants of accident patterns.

Moreover, the influence of peak-hour traffic on accident frequencies reflects the findings [12] that emphasized the role of dynamic road environments and pedestrian behavior in increasing bus accident risks. Rapid KL could consider adopting advanced video analytics or AI-based modeling, as suggested [12] to analyze pedestrian exposure and mitigate associated risks.

Additionally, the analysis of bus captain statuses (guilty versus not guilty) regarding accident costs showed that incidents, where captains were found guilty, had higher associated costs. This points to potential areas for improvement in driver training and monitoring and the introduction of stricter compliance and disciplinary measures. Segmenting bus driver risks based on age, experience, and historical crash data. Implementing such a segmentation approach could help Rapid KL enhance driver-specific interventions [13].

Several patterns emerged from the data, most notably the correlation between accident frequency, financial loss, and time of year. August and October were identified as high-risk months, driven primarily by external factors such as road conditions and increased passenger load. The polynomial trend line fitted to the data illustrated a general upward trend in losses mid-year, with peaks in August and October followed by a gradual decline towards the end of the year.

The categorization of accidents by involved parties revealed that vehicle-related incidents were the most frequent, contributing more than RM500,000 in losses, followed by road condition-related accidents. By concentrating efforts on high-risk routes, optimizing schedules during peak periods, enhancing bus fleet maintenance,

and improving driver awareness and reporting protocols, Rapid KL can significantly reduce both the frequency and severity of accidents while minimizing associated financial losses.

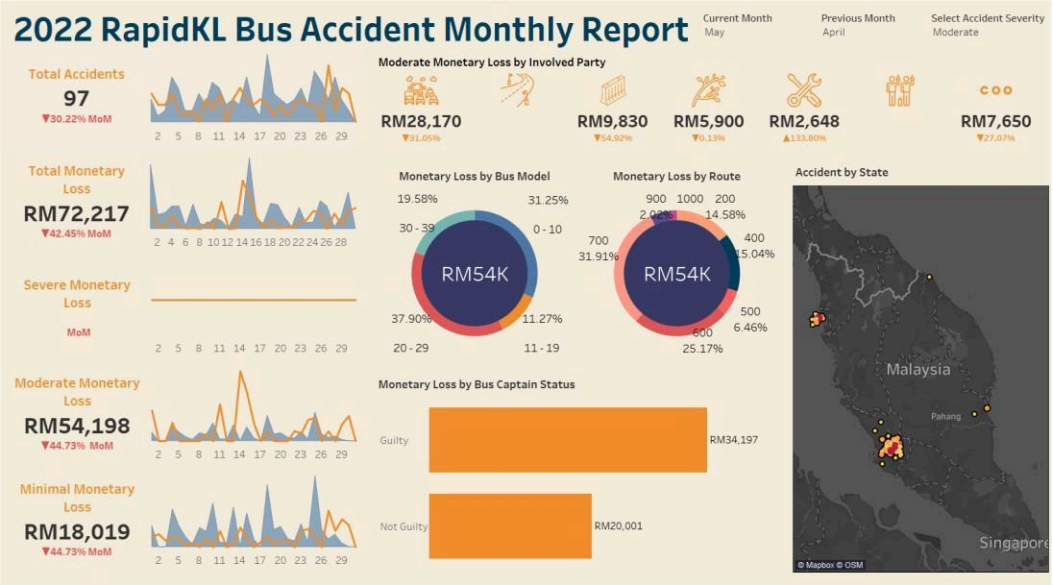

Figure 2. 2022 Rapid KL Bus Accident Monthly Report

4.2 Cost Analysis and Optimization

The analysis of the 2022 accident data revealed that vehicle-related incidents were the leading cause of accidents, accounting for 1,517 incidents and contributing RM1,046,364.13 to the overall financial losses. Collisions with immovable objects ranked as the second most common cause, with 267 incidents resulting in a substantial loss of RM1,204,920.17. Road conditions also played a significant role, leading to 95 accidents and causing RM148,299.69 in damages. Accidents due to nature-related factors, such as environmental hazards, amounted to 179 incidents with a total loss of RM144,013.63.

The data also highlighted other less frequent causes, including technical issues (54 incidents, RM82,563.53) and incidents involving passengers (65 incidents, RM18,514.91). Additionally, minor categories like "Other" (78 incidents, RM60,940.94) and a single "Natural Disaster" case (0 loss) were noted, and the month of August recorded the highest number of accidents (170 incidents) and financial loss (RM635,000), primarily driven by collisions with immovable objects and road conditions. These incidents were predominantly observed during evening peak hours (5 PM - 7 PM) when traffic congestion and driver fatigue were at their highest. These findings emphasize the need for targeted interventions during high- risk periods and improved maintenance of road infrastructure to mitigate similar incidents in the future.

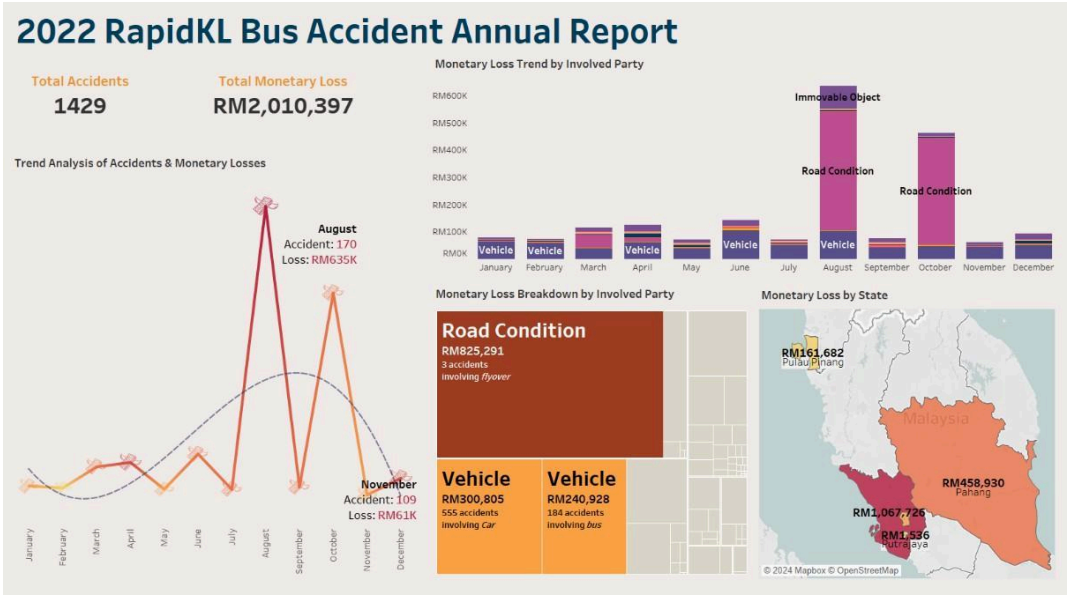

Figure 3. 2022 Rapid KL Bus Accident Annual Report

The financial impact of RM2,010,397 for the year highlights the pressing need for cost-effective strategies. Accident costs often escalate due to overlooked environmental factors like poor signage and bus stop crowding [12]. Rapid KL's emphasis on these areas could significantly reduce indirect costs and align with the findings.

Integrating traffic violation data into accident analytics is also essential for cost optimization [10]. Doing so for Rapid KL could strengthen preventive measures by identifying repeat offenders and correlating violations with accident hotspots.

4.3 Risk Assessment and Mitigation

The analysis of accident data and financial losses across different route groups reveals distinct patterns in the frequency of incidents and their associated monetary impact. The analysis highlights significant variations in accident counts and total monetary losses by categorizing routes into hundreds of hundreds. Route Group 400 stands out among the groups analyzed, with a substantial monetary loss of RM921,050 resulting from 144 accidents. Despite having a relatively moderate accident count compared to others, this group incurs the highest financial impact, indicating potentially severe incidents or routes with high-cost repairs.

On the other hand, Route Group 800 records the highest accident frequency, with 295 incidents leading to a total financial loss of RM280,962. This suggests that while Route Group 800 experiences more frequent accidents, the financial impact per incident is lower compared to Route Group 400. Route Group 700 follows closely with 264 accidents and a financial loss of RM253,317, contributing significantly to overall losses due to its high accident frequency. Similarly, Route Group 600, with 225 accidents, incurs a financial loss of RM244,972, closely resembling the impact of Route Group 700, albeit with fewer incidents. In contrast, Route Group 200 exhibits 278 accidents but a comparatively lower monetary loss of RM143,341, indicating that while the group has a high accident rate, the severity or cost per incident is relatively lower.

Route Group 900 represents a mid-range impact, with 125 accidents resulting in a financial loss of RM89,671. Route Group 500, with 63 accidents and a corresponding loss of RM40,990, is one of the minor impactful groups regarding accident count and financial losses. Additionally, a smaller group with other route IDs incurs a loss of RM24,529. This highlights that although some incidents are not categorized within the main groups, they contribute to the overall financial burden.

Analysis by route groups revealed disparities in accident frequencies and monetary losses. This mirrors findings that stressed the need for tailored interventions for high-risk routes, such as deploying IoT sensors and big data analytics for real-time monitoring and risk prediction [14].

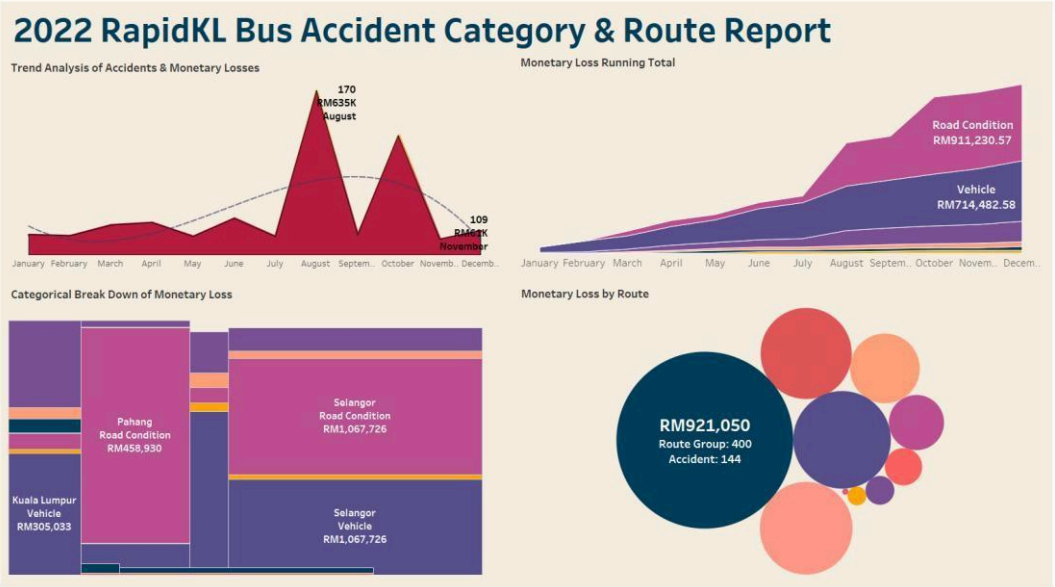

Figure 4. 2022 Rapid KL Bus Accident Category & Route Breakdown

Rapid KL's data insights resonate with broader trends highlighted in the literature, including the effectiveness of leveraging big data technologies and intelligent transportation systems to enhance safety and operational efficiency.

Future strategies for Rapid KL could include adopting advanced predictive models, integrating IoT-based safety measures, and using GIS for hotspot mapping.

5. Conclusion

The findings of this study provide critical insights into the patterns and contributing factors of Rapid KL bus accidents, aligning with broader trends identified in traffic safety research. However, certain limitations must be acknowledged. First, the dataset's reliance on internally reported accident data introduces potential biases, such as underreporting minor accidents or inconsistencies in data recording practices. Additionally, it did not incorporate external factors such as pedestrian behaviors, detailed weather conditions, or real-time traffic data, which are critical to understanding accident dynamics comprehensively. These omissions may limit the generalizability of the findings across different contexts.

Future research should address these gaps by integrating diverse data sources, including real-time traffic monitoring, environmental sensors, and external accident reports. Advanced methodologies, such as GIS-based spatial clustering or deep

learning for anomaly detection in bus trajectories, could further refine hotspot analysis and identify hidden risk patterns, as evidenced by techniques applied in the referenced studies. Moreover, leveraging big data analytics frameworks to analyze traffic violations and their correlation with accidents can provide actionable insights for law enforcement and traffic management.

Expanding the scope of analysis to include machine learning models for predictive analytics may offer more robust tools for accident prevention. For instance, employing neural network-based systems to predict accident severity or using spatiotemporal modeling for dynamic risk assessment can significantly enhance safety interventions. By adopting these advanced techniques and broader datasets, future studies can better inform policy decisions and operational strategies to optimize public transportation safety and efficiency.

Acknowledgements

This analysis was supported by Faculty of Artificial Intelligence, International Business School, Faculty of Social Sciences and Humanities, University Teknologi Malaysia

References

[1] S. Feng, Z. Li, Y. Ci, and G. Zhang, 'Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers', *Accid. Anal. Prev.*, vol. 86, pp. 29–39, Jan. 2016, doi: 10.1016/j.aap.2015.09.025.

[2] M. F. Musa, S. A. Hassan, and N. Mashros, 'The impact of roadway conditions towards accident severity on federal roads in Malaysia', *PLOS ONE*, vol. 15, no. 7, p. e0235564, Jul. 2020, doi: 10.1371/journal.pone.0235564.

[3] D. Henézi and Á. Winkler, 'The Role of Public Transport in Transport Safety and Public Safety', *Eurasia Proc. Sci. Technol. Eng. Math.*, vol. 23, pp. 505–512, Oct. 2023, doi: 10.55549/epstem.1374907.

[4] F. Vallet, M. Khouadjia, A. Amrani, and J. Pouzet, 'DESIGNING A DATA VISUALISATION AND ANALYSIS TOOL FOR SUPPORTING DECISION-MAKING WITH PUBLIC TRANSPORTATION NETWORK', *Proc. Des. Soc.*, vol. 1, pp. 1093–1102, Aug. 2021, doi: 10.1017/pds.2021.109.

[5] A. Clarinval and B. Dumas, 'Intra-City Traffic Data Visualization: A Systematic Literature Review', *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6298–6315, Jul. 2022, doi: 10.1109/TITS.2021.3092036.

[6] F. N. M. Nusa, S. Z. Ishak, R. Rusli, C. M. M. Isa, M. M. A. Manan, and S. Sulistyono, 'ROAD CRASH DATA VISUALISATION AND ANALYTICS USING TABLEAU FOR MOUNTAINOUS ROADWAY AREAS IN CAMERON HIGHLANDS, MALAYSIA', *Plan. Malays.*, vol. 21, Aug. 2023, doi: 10.21837/pm.v21i28.1314.

[7] V. Tandrayen-Ragoobur, 'The economic burden of road traffic accidents and injuries: A small island perspective', *Int. J. Transp. Sci. Technol.*, Mar. 2024, doi: 10.1016/j.ijtst.2024.03.002.

[8] N. Shaadan, M. I. K. A. Suhaimi, M. I. Hazmir, and E. N. Hamzah, 'Road accidents analytics with data visualization: a case study in Shah Alam Malaysia', *J. Phys. Conf. Ser.*, vol. 1988, no. 1, p. 012043, Jul. 2021, doi: 10.1088/1742-6596/1988/1/012043.

[9] M. Rabbani, M. A. Musarat, W. Alaloul, A. Maqsoom, H. Bukhari, and W. Rafiq, 'Road Traffic Accident Data Analysis and Its Visualization', *Civ. Eng. Archit.*, vol. 9, pp. 1603–1614, Aug. 2021, doi: 10.13189/cea.2021.090530.

[10] H.-C. Yang, M.-C. Chen, and I.-L. Lin, 'Discussion on Application of Big Data Analysis in Improving Traffic Safety: Database of Traffic Accidents and Violation Reports', in *2023 IEEE 3rd International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB)*, Apr. 2023, pp. 219–222. doi: 10.1109/ICEIB57887.2023.10170714.

[11] K. Alkaabi, 'Identification of hotspot areas for traffic accidents and analyzing drivers' behaviors and road accidents', *Transp. Res. Interdiscip. Perspect.*, vol. 22, p. 100929, Nov. 2023, doi: 10.1016/j.trip.2023.100929.

[12] B. P. Y. Loo, Z. Fan, T. Lian, and F. Zhang, 'Using computer vision and machine learning to identify bus safety risk factors', *Accid. Anal. Prev.*, vol. 185, p. 107017, Jun. 2023, doi: 10.1016/j.aap.2023.107017.

[13] T. Ding, H. Yin, Z. Li, X. He, L. Zheng, and J. Xi, 'Analysis of Risk Bus Driver Characteristics and Research on Risk Level Evaluation Methods for Bus Drivers', *IEEE Access*, vol. 12, pp. 171348–171367, 2024, doi: 10.1109/ACCESS.2024.3498936.

[14] M. K. Sengul, C. Tarhan, and V. Tecim, 'Application of Intelligent Transportation System Data using Big Data Technologies', in *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Sep. 2022, pp. 1–6. doi: 10.1109/ASYU56188.2022.9925457.