

“© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

ExTraVis: Exploration of Traffic Incidents Using Visual Interactive System

Joshua Zerafa*, Md Rafiqul Islam*, Muhammad Ashad Kabir†, Guandong Xu*

*Advanced Analytics Institute (AAI), University of Technology Sydney, Australia

†School of Computing and Mathematics, Charles Sturt University, NSW, Australia

{Joshua.Zerafa, mdrafiqul.islam-1}@student.uts.edu.au, akabir@csu.edu.au, guandong.xu@uts.edu.au

Abstract—The impact of road traffic incidents (e.g., road accidents, vehicle breakdowns) have become progressively worse over the years, being a major cause of many adverse issues such as serious injury, economic loss, and lifelong disabilities. Thus, it is essential to acknowledge these issues and proactively construct appropriate solutions to mitigate the impact of these issues in the future. This study outlines the history of traffic incident research and covers several unique solutions such as *Traffic Origins*, *Fervor* to traffic incident analysis. In this paper, we design a unique visualization system, *ExTraVis*, for incident data exploration and analysis that can be used to help traffic management controllers, aid to make decisions, and help them to understand how past incidents affected and where incidents may occur. The key features of this system are visual exploration and analysis to overcome the problems linked with road traffic incidents and to encourage future work and improvements. Finally, we demonstrate the effectiveness of *ExTraVis* by comparing with three different incident visualization dashboards.

Index Terms—Data visualization, traffic incident, incident management.

I. INTRODUCTION

Road traffic incidents such as road accidents and vehicle breakdowns are the cause of many fatalities, injuries, financial, and mental health problems amongst the people affected, as well as the nation as a whole [1], [2]. Consequently, it is important to be able to analyze these incidents and their extremity as an important step in mitigating or preventing the threatening complications that come with it [3]. Approximately 1.35 million people die each year as a result of road traffic crashes [4]. According to World Health Organisation (WHO), road traffic injuries are the leading cause of death for children and young adults aged 5 to 29 years old. Additionally, these types of injuries also cause considerable economic loss, costing most countries 3% of their gross domestic product [5].

As a result of this issue, an ongoing field of research is traffic incident analysis. While considering the topic of traffic incident analysis, there were several questions raised such as 1) Why is traffic incident analysis necessary?, 2) What effect will advancements in traffic incident analysis have in minimising the impact of these incidents?, 3) What are the current solutions to traffic incident analysis?, and 4) How can existing approaches to traffic incident analysis be improved upon to develop better solutions? Thus, with careful consideration of these questions, we can better understand the significance of this area of research and critically evaluate

the current research methods and contributions made to traffic incident analysis to identify future work and improvements [6].

The development of various solutions in the form of machine learning [3], [7], mathematical modeling [8], [9], and visualization system [5] are designed to analyse the likelihood of incidents occurring under certain spatio-temporal conditions [7], [10]–[12]. Many of these existing solutions (e.g., [13], [14]) utilise real-time traffic and incident data to accurately estimate the duration of an incident and clearance time. Machine learning models represent real-world processes through the notion of self-learning. In comparison, visualization system methodically organise traffic incident data in order to illustrate important trends and information that can be used to analyse where incidents may occur based on spatial and temporal conditions [15], [16].

In this paper, a visualization system named *ExTraVis* is designed using a traffic incident dataset containing all incidents which occurred in South Australia in 2019 [17]. Figure 1 illustrates an interactive dashboard consisting of organised traffic incident data for visual exploration and analysis. Although existing many visualization systems such as Incident by Borough (IB) [18], Fatality Analysis Reporting System (FARS) [19], and Incident Trend Dashboard (ITD) [20] are designed for data exploration and analysis of traffic incident, they contain a myriad of unnecessary features that leave the design overly complex and ineffective [2]. On the other hand, *ExTraVis* enables the user to quickly identify trends of traffic incidents in selected areas to effectively implement traffic management strategies to minimize incident records. The features included in the dashboard have been carefully selected to ensure users with all levels of experience and knowledge, and to effectively make use of the visualization system for learning, business decisions, and other interests.

The main contribution of this paper is to explore and analyse incident data visually where *ExTraVis* is comprised several key features such as 1) Free Text Search to quickly locate information, 2) Incident trend chart with incident forecasting, 3) Map for visualization of incident location and frequency, 4) Bar chart illustrating suburbs with the highest incident count, 5) Pie chart showing what part of the state these incidents occur, 6) Demonstrates the severity of these incidents, and 7) Chart showing frequency of various crash types.

The rest of the paper is organized as follows: In Section II, we review existing works on the significance of traffic features,

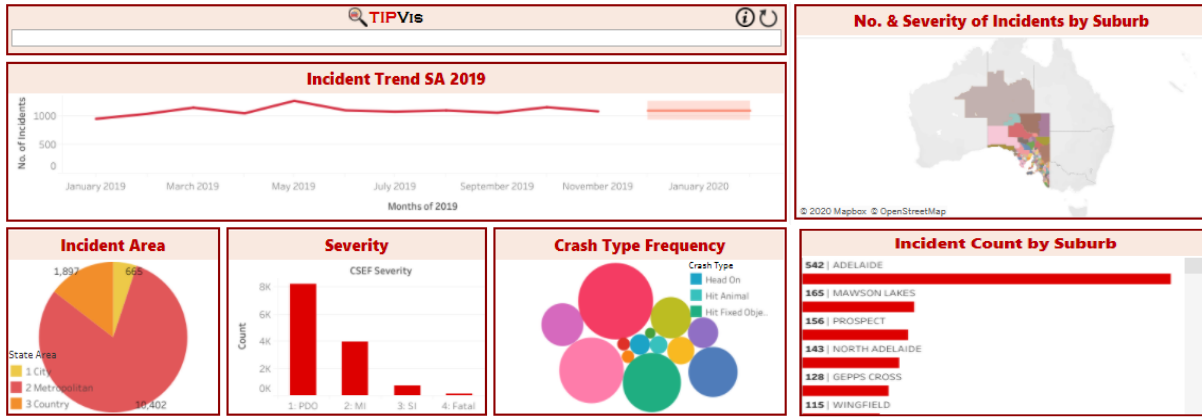


Fig. 1. Overview of ExTraVis dashboard for traffic incident prediction

various solutions, and identify limitations of current research. In Section V, we describe the design framework of our proposed system, and in Section VI, we provide result and analysis of the system. Finally, Section VIII concludes the paper with directions of future work.

II. RELATED WORK

In this section, we discuss the significance of real-time traffic data for traffic incident analysis, the existing approaches to analyse traffic incidents and their limitations.

A. Significance of Traffic Incident Analysis

Sun et al. [21] has investigated various spatiotemporal impacts of traffic incidents which are faced in road networks. They mainly considered four road network features such as Betweenness Centrality, weighted PageRank, Hub, and K-shell which would help in evaluation of incident impacts on urban traffic mobility. Javid and Javid [22] have proposed a series of robust regression based models which would help in proper estimation of travel time variability which is being caused by traffic incidents using various integrated data related to traffic, geometry, incidents caused and weather conditions. Kaddoura and Nagel [23] clearly depict the significance of real-world data for use of modelling as it accounts for more realistic cases. In their research, temporal and spatial were the two measurements of incident impact. The determination of a strong correlation between incident delay and two network features are the component of hazard-based models. Through their research, we are able to see the complications when using skewed data to evaluate model performance. Therefore, the existing research establishes that the location of incident sites and the functionality of an intersection in a network are associated with incident impact.

B. Existing solutions

Javid and Javid [22] developed a framework for estimating time variability produced by traffic incidents using urban traffic incident data. The study analyses the accuracy of the model over a period of two years in California's highway

system. Haule et al. [24] use Hazard Based models to analyze the factors relating to the impact of incident duration using a new performance measure and the traditional incident clearance duration. Anwar et al. [14] propose 'Traffic Origins', a visualization technique for supporting of traffic incident analysis. Pettet et al. [25] present a visual dashboard which enables emergency responders to analyse and manage spatio-temporal incidents including crime and traffic incidents. Hu et al. [26] propose a three-dimensional model-based vehicle tracking probabilistic model which applies neural network algorithms to data in the form of motion trajectories to learn activity patterns. The model predicts vehicle activity from matching each partial trajectory with the learned activity patterns.

C. Limitations

The existing research tends to include data from a limited number of datasets and does not use real traffic data. They do not provide a realistic representation of the accuracy and performance of the proposed solutions, as datasets which knowingly provide the greatest results have been selected. Furthermore, the machine learning (ML) models presented in this review face difficulties accounting for the influence of outliers in datasets in the final prediction of an incident's duration. The algorithm of ML models should be adjusted to properly account for outliers, possibly by introducing a weight to each data point to represent its significance [27]. However, when synthesising and evaluating the literature, it is clearly evident that there are many limitations and gaps in this area of research. An adequate solution to these limitations could provide a means to mitigate the problems that arise as a result of road traffic incidents [24]. For example, visual exploration and analysis solutions allow for a more simplistic process of trend identification and can effectively measure and evaluate performance to more accurately forecast traffic incidents.

Evidently, the outlined limitations embedded within existing research have inspired this study of data visualization for exploration and analysis. We believe that data visualization

TABLE I
DATASET DESCRIPTION

Features description	Values
Total number of Data	2300
Total number of attributes	33
Data type	Srting
Data history	South Australian

offers a range of benefits that cannot be acquired through machine learning approaches [1].

III. DATA DESCRIPTION AND PROCESSING

The dataset used in this paper was collected from the South Australian Government Data Directory [28]. The dataset consisting of historical data of incidents which occurred all throughout the state of South Australia in 2019. It contains various attributes describing details about each incident, which are outlined in Table I. The original dataset consisted of 33 attributes. In order to simplify the dashboard to ensure only the most significant data is used, data preprocessing consisted of identifying and eliminating less important and redundant attributes which offer no benefit to exploration and analysis. As part of data preprocessing, redundant fields that were not eliminated, were combined. For example, individual attributes ‘Month’ and ‘Year’ were combined to form a single attribute ‘Month & Year’. Additionally, several new calculated fields were created to keep track of entries in the Free-Text-Search using a free text filter as well as Boolean fields.

IV. SYSTEM DESIGN

In this section, we describe *ExTraVis* which is designed as a tool to visually explore and analyse traffic incident data. Similar to many web applications and dashboards [29], [30], our proposed dashboard is comprised three primary parts: back-end, gateway and front-end.

The **back-end part** of the framework contains the dataset that forms the basis of our dashboard. The dataset then undergoes filtering and transformation to restructure the data into a proper format for querying and analysis. The dataset is then connected to the storage components using data connectors namely the Data Engines, SQL, and MDX. These connectors enable the technology to manipulate the dataset for almost every action in our dashboard. Our storage component has an SQL server and file storage for storing the required codes for our dashboard.

Our back-end and front-end are connected via a **gateway** that processes the request received from the user side and gives the response to the front-end side.

In the **front-end part**, users can use both mouse and keyboard inputs to get a visual response. Our visual responses consist of a map, a bar chart, a line graph, and a pie chart. All our visual responses are dynamic, if the dataset is changed, our visual representation will change dynamically. Moreover, all the sections in the visual response are multiple coordinated, this means our dashboard has a link with various visual action/operation filtering.

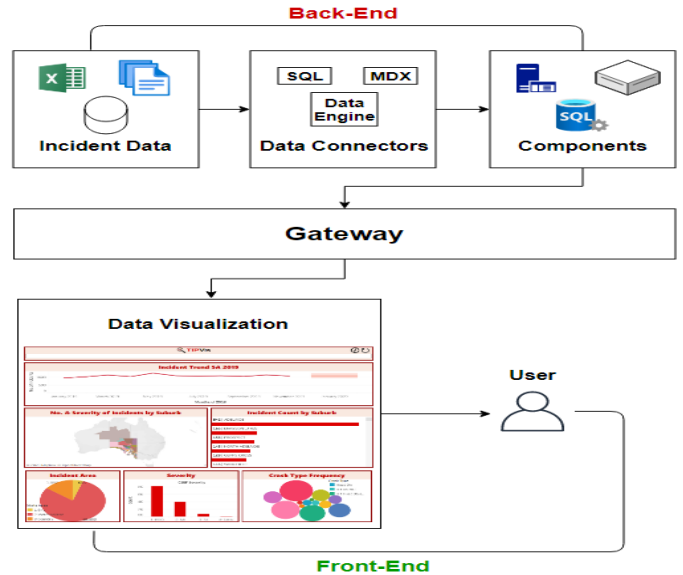


Fig. 2. The methodology of ExTraVis dashboard.

ExTraVis has implemented in Tableau (a powerful interactive visualization tool). Tableau is primarily chosen as it simplifies data analysis, being a faster and more effective approach in comparison to alternatives such as Microsoft Power BI or Qlik Sense [31].

V. DESCRIPTION OF EXTRAVIS

In this section, we briefly describe the design of *ExTraVis* and how each view and interaction feature supports the user analytic tasks. *ExTraVis* consists of seven views as Figure 1 presented in Section I. In the following, we dive deeper into the design choices of *ExTraVis* including an individual break down of each of the components, features and worksheets that make up the dashboard.

1. Free Text Search: The free-text-search (FTS) enables the user to quickly search through the data by Suburb, Month, Severity, Crash Type, and State Area. For example, if the user is seeking information on all incidents which occurred in Adelaide, they could locate and click on the city of Adelaide on the map, however, this is more time-consuming and degrades the user experience. An information button is included in the FTS to give a brief overview of what data types can be searched. Additionally, the reset button is used to return all parameters of the dashboard to their default settings.

2. Incident trend for forecasting: The incident trend line graph is significant in determining the likely number of incidents. This can be beneficial overtime to evaluate whether strategies or programs put in place to minimise traffic incidents within South Australia are effective or not. Hovering the cursor over a particular point on the graph will display whether the point is an estimate or real data, the month that has been selected, and the number of incidents which occurred. This worksheet is interconnected, hence when a point on the graph is selected, the other parts of the dashboard will illustrate

all information on the incidents which occurred during that particular month.

3. No & Severity of Incidents by Suburb: The visual map allows the user to zoom-in/zoom-out and scroll around to locate the various affected areas geographically. The inclusion of this feature further enables the user to effectively generate visualization of data for particular areas without having to determine the name for which to search. Additionally, this feature enables the user to simply visualize suburbs containing the most severe incidents to determine areas which require attention to improve traffic control.

4. Incident Count by Suburb: The incident count by suburb bar graph illustrates the name of each suburb throughout South Australia in order of the number of incidents that occurred.

5. Incident Incident Chart: The incident area chart classifies each incident into three separate area types: City, Metropolitan, and Country. The chart shows the number of incidents located in each type of area.

6. Incident Severity Graph: The incident severity graph categorises each incident into particular severity groups. These severity groups include: PDO: The incident resulted in property damage only, MI: The incident resulted in minimal injuries, SI: The incident resulted in serious injuries, and Fatal: The incident resulted in one or more fatalities.

7. Crash Type Frequency Chart: The crash type frequency chart illustrates the number of incidents which occurred from each crash type.

VI. RESULT AND DISCUSSION

From Table II, it is evident that the existing data visualization tools lack in Free Text Search (FTS) feature. However, FTS supports user queries by allowing users to search information to quickly locate and visually analyse data within the dashboard [29]. In developing this system, a ‘free text search’ parameter was defined to enable the several worksheets within the dashboard to essentially communicate and display data based on user search and selection. A ‘free text filter’ calculated field was created that referenced the FTS parameter. The filter tells the system to search for the field name from the dataset and display the data as a visual result in the dashboard. It also accounts for cases where the user searches specific keywords before typing the field name, e.g. ‘Show me incidents in Adelaide’.

According to Figure 1, the incident trend line graph demonstrates the number of incidents which occurred within South Australia in each month from January 2019 – November 2019 (represented by the dark-shaded red line) as well as a 3 month forecast for the number of incidents that should occur from December 2019 – February 2020 (represented by the greyed out red line). The number of incidents is conveyed along the Y-axis with the month in which they occur along the X-axis. The visual map illustrates the number of incidents which occurred across all the suburbs of South Australia in 2019. Particularly, selecting a individual suburb create an outline around it and grey out all other parts as well as display a text-box which

TABLE II
FEATURE COMPARISON

Features of dashboard	IB	FARS	ITD	<i>ExTraVis</i>
Free text search	✗	✗	✗	✓
Incident trend graph	✓	✗	✓	✓
Visualization map	✓	✗	✗	✓
Incident count by area	✓	✓	✗	✓
Area incidents occurred	✗	✗	✓	✓
Severity of incidents	✗	✓	✗	✓
Cause of crash	✗	✓	✓	✓

Notes: (✓) indicates the presence of feature selections and (✗) indicates that it does not visualize the feature outcome.

shows the number of incidents in that suburb in 2019 and the greatest severity.

The incident count by suburb worksheet is interconnected, hence when a point on the graph is selected, the other parts of the dashboard display information on the incidents which are located in that particular suburb. Moreover, the incident area chart shows the number of incidents located in each type of area. Hovering over a particular area displays a text-box which shows the area type and the number of incidents. When the user selects a particular area type, all the other worksheets of the dashboard display information on incidents which are located in that selected area, and unrelated data is greyed out. Finally, the crash type frequency chart opens a text-box which shows the type of crash and the number of incidents. When the user selects a particular crash type, other worksheets are adjusted to display data based on incidents which occurred from the selected crash type and unrelated data is greyed out.

In summary, we deduce the following key insights:

- *ExTraVis*, illustrates important trends and information that can be used to understand how past incidents affected and where incidents may occur.
- *ExTraVis*, allows Traffic incident controllers (TIC) to quickly search through the data by Suburb, Month, Severity, Crash Type, and State Area for traffic incidents analysis.
- It is observed that, the most risky traffic incident suburbs are Adelaide, Mawson Lakes, and Noeth Adelaide. Thus, TIC should pay more attention in these areas.

VII. EVALUATION

In order to evaluate the potential utility of our dashboard, in Table 2, *ExTraVis* is compared with three different incident visualization dashboards: Incident by Borough (IB), Fatality Analysis Reporting System (FARS), and Incident Trend Dashboard (ITD). The first, IB, captures data about a range of incident types including onboard injuries, slip/trip/fall, personal injury, collision incident, assault, safety critical failure, activity incident event, and vandalism hooliganism. Similarly to *ExTraVis*, the dashboard conveys a list of areas in which these incidents occur based on incident count and consists of a tree map for visualization. IRS effectively captures and categorises data on incidents involving customers, injury, spills, transport, security, equipment, complaints, and divergence. Similarly to our dashboard, IRS displays the cause and severity of these

incidents. Lastly, ITD captures data on incidents resulting from IT issues. Incidents include Network & Connectivity, Account & Access, Application & Tools, Application & Software, Alert & Monitoring, and External & Supplier. The dashboard includes an incident trend graph to illustrate the time period over which these incidents occur [27].

VIII. CONCLUSION AND FUTURE WORK

In this paper, we present *ExTraVis*, an interactive visualization system to provide a means of identifying important trends in the field of traffic incident analysis and encourage the implementation of suitable programs as well as the necessary steps to ensure better safety on our roads. We discovered and explored several interesting findings based on a similar foundation of traffic incidents analysis which could effectively benefit from a solution in the form of a visualization tool to identify patterns and trends from which conclusions can be drawn.

In future, the proposed data visualization tool should be combined with an NLP-based data preprocessing method to optimize the data being analysed, which can then be used to power a deep-learning algorithm to effectively predict trends of future incidents. Merging the interactive visual dashboard with an NLP data preprocessing method and a predictive deep-learning algorithm will enable the model to continually make improvements and presumably outperform existing machine learning models. Additionally in future works, *ExTraVis* should be extended to other fields of research by generalizing the dashboard to work with various types of incidents. This can be done by broadening the features of the dashboard to account for general incidents as opposed to traffic incidents in particular and enable the user to create custom parameters to suit their line of work.

REFERENCES

- [1] K. Fu, T. Ji, L. Zhao, and C.-T. Lu, "Titan: A spatiotemporal feature learning framework for traffic incident duration prediction," in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2019, pp. 329–338.
- [2] R. Li, F. C. Pereira, and M. E. Ben-Akiva, "Overview of traffic incident duration analysis and prediction," *European transport research review*, vol. 10, no. 2, p. 22, 2018.
- [3] L.-L. Wang, H. Y. Ngan, and N. H. Yung, "Automatic incident classification for large-scale traffic data by adaptive boosting svm," *Information Sciences*, vol. 467, pp. 59–73, 2018.
- [4] Y. Zou, X. Ye, K. Henrickson, J. Tang, and Y. Wang, "Jointly analyzing freeway traffic incident clearance and response time using a copula-based approach," *Transportation research part C: emerging technologies*, vol. 86, pp. 171–182, 2018.
- [5] K. Wongsuphasawat, M. L. Pack, D. Filippova, M. VanDaniker, and A. Olea, "Visual analytics for transportation incident data sets," *Transportation research record*, vol. 2138, no. 1, pp. 135–145, 2009.
- [6] I. Lana, J. Del Ser, M. Velez, and E. I. Vlahogianni, "Road traffic forecasting: Recent advances and new challenges," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 2, pp. 93–109, 2018.
- [7] R. Yu, Y. Li, C. Shahabi, U. Demiryurek, and Y. Liu, "Deep learning: A generic approach for extreme condition traffic forecasting," in *Proceedings of the 2017 SIAM international Conference on Data Mining*, SIAM, 2017, pp. 777–785.
- [8] A. MacEachren, D. Xiping, F. Hardisty, D. Guo, and G. Lengerich, "Exploring high-d spaces with multiform matrices and small multiples," in *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No. 03TH8714)*. IEEE, 2003, pp. 31–38.
- [9] M. A. A. Kuhail and S. Lauesen, "Uvis: A formula-based end-user tool for data visualization," *Ieee Access*, vol. 8, pp. 110 264–110 278, 2020.
- [10] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 606–616, 2011.
- [11] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *Journal of Transportation Engineering*, vol. 117, no. 2, pp. 178–188, 1991.
- [12] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *arXiv preprint arXiv:1709.04875*, 2017.
- [13] S. Boyles, D. Fajardo, and S. T. Waller, "A naive bayesian classifier for incident duration prediction," in *86th Annual Meeting of the Transportation Research Board, Washington, DC*. Citeseer, 2007.
- [14] A. Anwar, T. Nagel, and C. Ratti, "Traffic origins: A simple visualization technique to support traffic incident analysis," in *2014 IEEE Pacific Visualization Symposium*. IEEE, 2014, pp. 316–319.
- [15] C. Lee, Y. Kim, S. M. Jin, D. Kim, R. Maciejewski, D. Ebert, and S. Ko, "A visual analytics system for exploring, monitoring, and forecasting road traffic congestion," *IEEE Transactions on Visualization and Computer Graphics*, 2019.
- [16] M. Vassell, O. Apperson, P. Callyam, J. Gillis, and S. Ahmad, "Intelligent dashboard for augmented reality based incident command response coordination," in *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2016, pp. 976–979.
- [17] *Exploration of Traffic Incidents*. [Online]. Available: <https://public.tableau.com/IncidentForecastingDashboard>
- [18] *Incident by Borough*. [Online]. Available: <https://public.tableau.com/incidents by borough>
- [19] *Fatality Analysis Reporting System (FARS)*. [Online]. Available: [https://public.tableau.com/Fatality Analysis Reporting System \(FARS\)](https://public.tableau.com/Fatality Analysis Reporting System (FARS))
- [20] *Incident Trend Dashboard*. [Online]. Available: <https://public.tableau.com/Incident-Trend-Dashboard>
- [21] C. Sun, X. Pei, J. Hao, Y. Wang, Z. Zhang, and S. Wong, "Role of road network features in the evaluation of incident impacts on urban traffic mobility," *Transportation research part B: methodological*, vol. 117, pp. 101–116, 2018.
- [22] R. J. Javid and R. J. Javid, "A framework for travel time variability analysis using urban traffic incident data," *IATSS research*, vol. 42, no. 1, pp. 30–38, 2018.
- [23] I. Kaddoura and K. Nagel, "Using real-world traffic incident data in transport modeling," *Procedia computer science*, vol. 130, pp. 880–885, 2018.
- [24] H. J. Haule, T. Sando, R. Lentz, C.-H. Chuan, and P. Alluri, "Evaluating the impact and clearance duration of freeway incidents," *International journal of transportation science and technology*, vol. 8, no. 1, pp. 13–24, 2019.
- [25] G. Pettet, A. Mukhopadhyay, C. Samal, A. Dubey, and Y. Vorobeychik, "Incident management and analysis dashboard for fire departments: Iccps demo," in *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, 2019, pp. 336–337.
- [26] W. Hu, X. Xiao, D. Xie, T. Tan, and S. Maybank, "Traffic accident prediction using 3-d model-based vehicle tracking," *IEEE transactions on vehicular technology*, vol. 53, no. 3, pp. 677–694, 2004.
- [27] M. R. Islam, S. Liu, X. Wang, and G. Xu, "Deep learning for misinformation detection on online social networks: a survey and new perspectives," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–20, 2020.
- [28] *Dataset*. [Online]. Available: <https://data.gov.au/organisations/org-sa-b52dfd1f-cb00-4b0a-a6b8-0b3662066376>
- [29] M. R. Islam, S. Liu, I. Razzak, M. A. Kabir, X. Wang, and G. Xu, "Mhivis: Visual analytics for exploring mental illness of policyholder's in life insurance industry," in *2020 7th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESCC)*. IEEE, 2020.
- [30] M. T. Islam, M. R. Islam, S. Akter, and M. Kawser, "Designing dashboard for exploring tourist hotspots in bangladesh," in *The 23th International Conference on Computer and Information Technology (ICCIT-2020)*. IEEE, 2020.
- [31] *Tableau*. [Online]. Available: <https://www.tableau.com/products/new-features/ask-data>