

# Road Traffic Accident Data Analysis and Its Visualization

Muhammad Babar Ali Rabbani<sup>1</sup>, Muhammad Ali Musarat<sup>2</sup>, Wesam Salah Alaloul<sup>2,\*</sup>, Ahsen Maqsoom<sup>3</sup>, Hamna Bukhari<sup>4</sup>, Waqas Rafiq<sup>2,3</sup>

<sup>1</sup>Department of Civil and Engineering, Sarhad University of Science & Information Technology Peshawar, Pakistan

<sup>2</sup>Department of Civil and Environmental Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 32610 Tronoh, Perak, Malaysia

<sup>3</sup>Department of Civil Engineering, COMSATS University Islamabad Wah Campus, Wah Cantt, Pakistan

<sup>4</sup>NIT-SCEE, National University of Sciences and Technology, Islamabad, Pakistan

*Received May 4, 2021; Revised June 4, 2021; Accepted July 19, 2021*

## *Cite This Paper in the following Citation Styles*

*(a): [1] Muhammad Babar Ali Rabbani, Muhammad Ali Musarat, Wesam Salah Alaloul, Ahsen Maqsoom, Hamna Bukhari, Waqas Rafiq, "Road Traffic Accident Data Analysis and Its Visualization" Civil Engineering and Architecture, Vol. 9, No. 5, pp. 1603-1614, 2021. DOI: 10.13189/cea.2021.090530.*

*(b): Muhammad Babar Ali Rabbani, Muhammad Ali Musarat, Wesam Salah Alaloul, Ahsen Maqsoom, Hamna Bukhari, Waqas Rafiq (2021). Road Traffic Accident Data Analysis and Its Visualization. Civil Engineering and Architecture, 9(5), 1603-1614. DOI: 10.13189/cea.2021.090530.*

Copyright©2021 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** Vehicle accidents take human life all over the world particularly in developing countries like Pakistan. It is estimated that 1.2 million people lose their lives in road accidents every year. Apart from this, 20 to 50 million are injured on a yearly basis. This annual increase in the traffic accidents trend is alarming. To bring improvement in the current road network system, the specialists need to analyze the historical data of road crashes of an area. This research aims to use the visualization technique to have a better understanding of the accident data. This study uses the data of Peshawar, Pakistan, where the raw data were first organized, filtered, pre-processed and finally, visualization was performed to construct a systemic and homogenous data model. Various infographics were produced with the help of different software interface and visualization options. It was revealed that most of the accidents occur in the daytime and with those people who do not have enough traffic education. The 30 to 45 years age group was more active in causing the accidents. Therefore, the behaviour of this age group of drivers needs further investigation. This study will be useful for concerned authorities in devising an efficient mechanism to alleviate road accident cases.

**Keywords** Road Traffic Accidents (RTAs), Traffic Safety, Traffic Analysis, Visualization

---

## 1. Introduction

The presence of traffic is one of the unwanted gifts of urbanization [1-3]. Transportation is a type of infrastructure facility constructed by the government for its citizen for a quick and uninterrupted movement. It helps connect people, in commercial movements and also ensures the ease of accessibility [4-7]. However, failure to abide by road regulations and poor road conditions often result in crashes. These crashes not only cause destruction to the community and other commodities involved but also cost human lives [8]. Road traffic accidents (RTAs), which may occur anytime at any place, has become a major reason for many deaths all over the world. The accident may occur due to many contributing factors, which depends on the time of the day, peak hour of the traffic, sex, age, geometric design of the road, type of vehicle, environmental conditions and vehicle occupancy [9-11].

Accidents have become an existential threat to human life [12]. Road collisions are the second important factor of death for people aging between 5 to 29 years and it is the third leading cause for the age group between 30-44

years [13]. With the advancement of the motorized industry, vehicle density has increased worldwide. The traffic intensity in many developing countries have increased considerably, however, the infrastructure remained the same, which put a strain on the obsolete road system. Hence, the developing countries are affected more by RTAs which is the major cause of death in these countries [14]. Road traffic crashes are the 8th leading cause of death of all ages [15]. There are 50 million injuries per year due to accidents, which are more than the total population of the two largest cities of the world combined i.e. Delhi and Tokyo [16]. In 2004, it was reported that the global traffic accident fatalities were 1.2 million deaths per year. Furthermore, it was also predicted that traffic injuries will surge to 65% in the next two decades if no mechanism is devised to combat accident prevention [17]. It has been estimated that around 1.35 million road users are killed worldwide, which estimated as 3700 fatalities per day [15]. Road accidents not only result in permanent damage to the affected road users but also lead to an economic burden like emergency services, property damage, legal and court expenses, insurance costs and workplace damages [18]. It has been estimated that road injuries cost USD 65 million all over the world and USD 518 million in low- and middle-income countries [19, 20].

The number of deaths from 1990 to 2017 due to traffic accidents worldwide was recorded. The age group of 15-49 years are the working class and are the persistent road users, therefore their numbers are the greatest among all the other age groups in the same year. In 2017, 49,068 children under the age of 5, 62,412 children between the age of 5 to 14 years, 669,058 people belonging the age group 15 – 49 years, 311,714 people lying in the age group of 50 – 69 years and 150,815 people above the age of 70 years lost their lives in the traffic accident cases, which accounts for a total of 1.24 million lives [21]. In comparison, the lives lost due to traffic accidents were 1.15 million annually. Out of these 3% lives lost were cyclists, 28% were bike riders and 23% were pedestrians who lost lives to the RTAs. Within every 24 seconds, a person dies due to the RTAs. In low to middle-income countries, despite having the lowest percentage of the vehicle as compared to high-income countries, the situation of RTAs is even worse [15]. These low to middle-income countries combined constitute 93% of the lives lost due to RTAs [22]. In 2019, RTAs was regarded as the 7<sup>th</sup> cause of most deaths in lower-income countries [23].

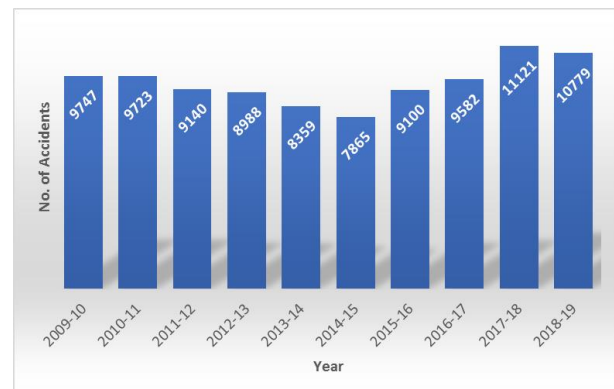
In the South Asian region, the highest number of accidents was recorded in India and stands the 60<sup>th</sup> country with the highest number of accidents reported in 2018. Similarly, China was ranked 89<sup>th</sup> in traffic accidents globally but in comparison to its neighbouring countries, China stands the 2<sup>nd</sup> in the highest number of accidents being reported. Pakistan stands 3<sup>rd</sup> in the South Asian

region with the highest number of accidents to happen. A description of South Asian countries traffic accidents is shown in Table 1.

**Table 1.** South Asian countries traffic accident statistics

| Country     | Number of accidents | World Rank in accidents | Source |
|-------------|---------------------|-------------------------|--------|
| Pakistan    | 30,046              | 95                      | [24]   |
| India       | 299,091             | 60                      | [25]   |
| China       | 284,986             | 89                      | [26]   |
| Iran        | 21,831              | 42                      | [27]   |
| Tajikistan  | 1,486               | 74                      | [28]   |
| Afghanistan | 5,253               | 90                      | [29]   |

According to World Health Organization (WHO), it was reported that the death rate from traffic accidents was 17.12% per 100,000 population, where, Pakistan ranked as the 95<sup>th</sup> country in accidents all over the world [30]. In Pakistan, the lowest RTAs were reported in the year 2014-15 while the highest RTAs were reported in the year 2017-18. A slight decreasing trend was observed from 2009 to 2015, but from 2015 to 2017, the RTAs started increasing with an average of 500 accidents annually [31]. The following Fig. 1 indicates the RTAs count.



**Figure 1.** RTAs in Pakistan from 2009 – 2019

A so-called solution that is unable to solve the problem is not even a solution. Similar is the case with having the RTAs statistics. No one could crack the code of statistical data until it is understood well enough, which can be solved by visualizing the data [32]. The reason is there are a lot of variables that are contributing to the event which cannot be predicted precisely. The need for data visualization is helpful in a way that it can open doors for the solution by understanding the data efficiently [33]. The RTAs data can play an influential role for the decision-makers, urban planners, transportation engineers and law enforcing agencies to implement the novelty ideas and vigorous solutions that can tackle the flaws in the current road system using the historical crash data [34].

The fundamental aim of this study is to use technology to help visualize the professionals the statistical data in-depth and become one step closer to interpret a lengthy

list of data. In other words, it is easier to look into graph and charts to extract data rather than looking into a segregated form of data. Without a proper understanding of the nature of the data, the information interpreted from the data is greatly jeopardized. This study puts into use the python coding environment to visualize the RTAs data. It uses the Python programming language software, which was first invented by Guido van Rossum in 1990. [35]. The added advantage of using Python is that its library is open source and one can easily import them and work on the project, which is termed as “batteries included” [36]. The libraries are placed in the Python Package Index (PyPI) that has libraries over 200,000 packages [37]. Some of the examples of the libraries are TensorFlow, Scikit-Learn, Pandas, Pytorch and various others [38]. Pandas and Scikit learn library was used in this study, which stands for “Python Data Analysis Library”. It is easier to work on this library because it can work on .csv, tsv or sql files and there is no need for a long list of coding dictionaries [39].

Whilst the world is moving into a digital age, one cannot rely on reading logs merely adding the data every day. At this pace, analyzing and understanding the data manually becomes a tiring and challenging task. Therefore, the need for visualizing the data using data science is of utmost importance as visualizing the historical accident data reveal the concealed facts. With pictorial illustration, it can be easily understood by the professionals and the traffic law enforcing agencies. Consequently, it will be helpful for the professional and decision-makers to counter the most contributing factors for the accident. Hence, the use of visualization is the need of the hour as it gives a detailed insight into the traffic accident which changes with time. This paper utilizes the historical traffic accident to know the contributing factors due to which accidents occur. It will be useful for planning engineers and decision-makers and will also help in traffic management and emergency service providers. The authorities can devise an effective master plan for the mitigation of accidental cases.

## 2. Literature Review

Visualization comes in many shapes like a pie chart, bar chart, scatter plot, heat map or in a combination of all these charts or 3D plot depending on the nature of the data. Some advanced types of charts have evolved into a pyramid, statistical and contours charts but there is a drawback of organizing data into the charts, which is the requirement of sufficient space for large data [40, 41]. The visualization was performed on a time series based on United Kingdom road traffic accident data that helped the researchers to interpret the important information from the time series and forward it to the concerned agencies for better understanding [42]. A study was conducted on the

accident data taken from the traffic department of the United Kingdom (UK), which was visualized using Tableau software. It was stressed that the lengthy data should be visualized to have an in-depth understanding [43]. To assess the reasons for train derailment incidents, a Google Earth (GE) environment was put into application based on the real-life derailment incident data that took place in Hoxie, Casselton and Graettinger. A six degree of freedom visualization was performed based on the gross vehicle movement using 3-dimensional animation. The visualization helped to understand the topography and imagery of rails effectively [44]. The traffic data of Nepal was put into use to visualize the traffic network. For this purpose, the network file was analyzed into Log stash and finally, another working environment called Kibana was used for visualization. It was concluded that working with data science saves time and is cost-effective [45]. In another study, the traffic data was collected using location-based, activity-based and device-based data collectors. The study showed that there were many shapes of data visualization and also stated the steps that could be followed and make the visualization much easier. Though the visualization is performed offline its scope can be extended to online and live visual analytics obtained from the social media, which can be used in context with the Intelligent Transportation Systems (ITS) [46]. The Portugal traffic congestion data was analyzed using visualization. It was stressed that visualization plays a vital role in traffic modelling, Spatio-temporal dynamics and urban mobility tracks. This study used heat maps and space-time cubes for the visualization shapes [47]. The Khartoum state of Sudan has a major traffic problem. The RTAs data was gathered from several sources. The visualization of the data revealed that most accidents were reported in the monsoon season. Above all, young people were more prone to RTAs as compared to elderly people [48]. The RTAs visualization was performed using the R programming package. The data used for analysis belonged to the Tashkent city RTAs. The accidents were classified into Type I (collisions) and Type II (collisions with pedestrians). After visualizing it was found that both types of accidents happened at the same magnitude over the analysis period. The research findings were forwarded to the concerned traffic wardens and police officers for their effective planning [49].

## 3. Methodology

The visualization performed in this study was conducted by collecting data and then performing data analysis to make it homogenous for visualization purposes. This study conducted visualization by showing the number of fatal accidents, injury accidents and non-injury/non-fatal. Then, the number of accidents with respect to the most accident location was plotted in the pie

chart. Finally, the ArcGIS environment was used to plot the accidents on the Hayatabad, Peshawar map using its coordinates.

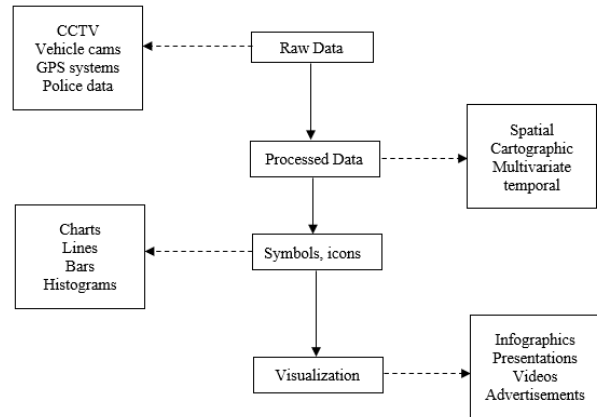
**3.1. Data Collection**

The historical data of RTAs was collected for the Hayatabad area, which lies in the Peshawar region situated in the North West of Khyber Pakhtunkhwa province of Pakistan. The data was collected from the local Police station, which keeps records of accidents taking place every day. The data ranges from 2009 to 2020 containing 1819 vehicle accident cases. The selected data consists of all two incidents i.e. death and injuries, also containing a total number of accidents that took place in a single day. The accident data contains driver information, accident information and the site of the accidents. The dataset was analyzed using comma-separated values (.csv) format. Each accident was identified by the drivers' age, education, gender, presence of license, vehicle type and the location of the accident. The data description is provided in Table 2.

**3.2. Data Analysis**

The visualization consists of four steps namely: Raw data, processed data, visual symbols and visualization. In the first step, the data was collected from various sources like official local department statistics, GPS data or incident logs. In the preprocessing step, the data was

converted into temporal properties. In the third step, visual symbols like bar and line charts were made. The final step includes the formation of infographics of maps or coloured image or the combination as it is user-dependent. This research flowchart is provided in Fig. 2.



**Figure 2.** Research Flowchart

In this study, various visualization methods were used. Infographic was made to create an overview of all the accidents that happened from 2009 to 2020. The description was also made part of the infographics that included the age group accidents, accidents involving gender, time of the day, location of the accident and class of vehicles.

**Table 2.** Data description of the Hayatabad RTAs

| ID | time  | No of deaths | No. of injuries | Total no. of accident | Education level | Age | Gender | License | Vehicle type | Location                |
|----|-------|--------------|-----------------|-----------------------|-----------------|-----|--------|---------|--------------|-------------------------|
| 1  | night | 0            | 2               | 1                     | intermediate    | 22  | male   | no      | local        | RMI Hospital            |
| 2  | day   | 1            | 1               | 2                     | uneducated      | 44  | male   | yes     | commercial   | Sunday bazar            |
| 3  | night | 0            | 1               | 1                     | uneducated      | 35  | male   | no      | local        | H-1, street 3           |
| 4  | day   | 1            | 3               | 1                     | intermediate    | 19  | female | yes     | private      | Turangzai Market        |
| 5  | day   | 0            | 4               | 3                     | uneducated      | 54  | male   | no      | commercial   | Zarghoni Masjid         |
| 6  | night | 0            | 1               | 1                     | intermediate    | 22  | male   | no      | local        | ringroad flyover        |
| 7  | day   | 0            | 2               | 1                     | postgraduate    | 39  | male   | yes     | private      | PGMI centre, Ph-5       |
| 8  | night | 1            | 1               | 1                     | intermediate    | 34  | male   | no      | commercial   | Tatara roundabout       |
| 9  | day   | 0            | 1               | 2                     | intermediate    | 28  | female | yes     | private      | Jannat khaton park      |
| 10 | night | 2            | 3               | 1                     | uneducated      | 40  | male   | no      | private      | ringroad flyover        |
| 11 | day   | 0            | 2               | 1                     | intermediate    | 24  | male   | yes     | local        | Bagh-e-Naran roundabout |

### 4. Results and Discussion

The visualization performed revealed many interesting facts, which were hidden in the statistics and users were unable to identify them. The following Fig. 3 illustrates the fatal accident time, literacy of driver, gender, vehicle type and deaths per year. Most fatal accidents happened in the daytime ranging to 76 accidents. While most fatal accidents recorded were in the year 2020.

Similarly, the injury accident, as shown in Fig. 4, was also created to assess the associated data. Mostly, the uneducated class of drivers went into injury accidents. The male drivers were mostly responsible for the injury accidents and it ranged to 431 injury accidents. Moreover, the injury accidents of local transport vehicles were more involved than any other vehicle types. About 344 injury accidents consisted of local transport vehicles. Finally, the year 2020 remained the year with the most injury accidents having numbers of 189.

Furthermore, the non-fatal/non-injury accidents that resulted in damage to infrastructure and no human loss accidents were also plotted with the associated data as shown in Fig. 5. The non-fatal/non-injury accidents happened mostly in the daytime having 944 accidents from the period of 2009 to 2020. The male gender was mostly involved in this type of accident. The local transport vehicles accidents of non-fatal/non-injury accidents were greater than any other vehicle type having 891 accidents.

The overall number of accidents was 1819 from 2009 till 2020. This pie chart, as shown in Fig. 6, shows the major locations where the highest numbers of accidents happened. It also demonstrates the hotspots that are present in the Hayatabad area, which is incomprehensible by looking into the data directly. The dangerous locations for the most frequent accident locations are known as hotspots. The percentage of each location contributing to the RTAs (both fatal and injuries) was also highlighted.

| S.No | Comparison               | Description   | Statistics with maximum 100 number scale |
|------|--------------------------|---------------|--|
| 1    | Fatal vs Literacy Level  | Uneducated    | 40                                       |
|      |                          | Matric        | 12                                       |
|      |                          | Intermediate  | 6  |
|      |                          | undergraduate | 26                                       |
|      |                          | Postgraduate  | 33                                       |
| 2    | Fatal vs Time of the day | Day           | 76                                       |
|      |                          | Night         | 69                                       |
| 3    | Fatal vs Gender          | Male          | 50                                       |
|      |                          | Female        | 7  |
| 4    | Fatal vs Vehicle type    | Private       | 30                                       |
|      |                          | Local         | 18                                       |
|      |                          | Commercial    | 9  |
| 5    | No. of deaths per year   | 2009          | 10                                       |
|      |                          | 2010          | 17                                       |
|      |                          | 2011          | 17                                       |
|      |                          | 2012          | 11                                       |
|      |                          | 2013          | 16                                       |
|      |                          | 2014          | 17                                       |
|      |                          | 2015          | 11                                       |
|      |                          | 2016          | 19                                       |
|      |                          | 2017          | 22                                       |
|      |                          | 2018          | 24                                       |
|      |                          | 2019          | 27                                       |
|      |                          | 2020          | 31                                       |

Figure 3. Fatal accidents

| S.No | Comparison                         | Description   | Statistics with maximum 400 number scale |
|------|------------------------------------|---------------|--|
| 1    | Injury accidents vs Literacy Level | Uneducated    | 199                                      |
|      |                                    | Matric        | 72                                       |
|      |                                    | Intermediate  | 79                                       |
|      |                                    | undergraduate | 93                                       |
|      |                                    | Postgraduate  | 34                                       |
| 2    | Injury vs Time of the day          | Day           | 320                                      |
|      |                                    | Night         | 157                                      |
| 3    | Injury vs Gender                   | Male          | 431                                      |
|      |                                    | Female        | 104                                      |
| 4    | Injury vs Vehicle type             | Private       | 280                                      |
|      |                                    | Local         | 344                                      |
|      |                                    | Commercial    | 203                                      |
| 5    | Injuries in accident per year      | 2009          | 79                                       |
|      |                                    | 2010          | 88                                       |
|      |                                    | 2011          | 94                                       |
|      |                                    | 2012          | 81                                       |
|      |                                    | 2013          | 84                                       |
|      |                                    | 2014          | 101                                      |
|      |                                    | 2015          | 110                                      |
|      |                                    | 2016          | 120                                      |
|      |                                    | 2017          | 144                                      |
|      |                                    | 2018          | 152                                      |
|      |                                    | 2019          | 166                                      |
| 2020 | 189                                |               |  |

Figure 4. Injury accidents

| S.No | Comparison                             | Description   | Statistics with maximum 1000 number scale |
|------|--|---------------|---|
| 1    | Non-injury accident vs Literacy Level  | Uneducated    | 40  |
|      |  | Matric        | 12  |
|      |  | Intermediate  | 6   |
|      |  | undergraduate | 26  |
|      |  | Postgraduate  | 33  |
| 2    | Non-injury accident vs Time of the day | Day           | 944                                       |
|      |  | Night         | 622                                       |
| 3    | Non-injury accident vs Gender          | Male          | 994                                       |
|      |  | Female        | 321                                       |
| 4    | Non-injury accident vs Vehicle type    | Private       | 780                                       |
|      |  | Local         | 891                                       |
|      |  | Commercial    | 400                                       |
| 5    | Non-injury accidents per year          | 2009          | 31  |
|      |  | 2010          | 47  |
|      |  | 2011          | 44  |
|      |  | 2012          | 51  |
|      |  | 2013          | 59  |
|      |  | 2014          | 66  |
|      |  | 2015          | 79  |
|      |  | 2016          | 72  |
|      |  | 2017          | 88  |
|      |  | 2018          | 104                                       |
|      |  | 2019          | 112                                       |
| 2020 | 138                                    |               |   |

Figure 5. Non-fatal/non-Injury accidents

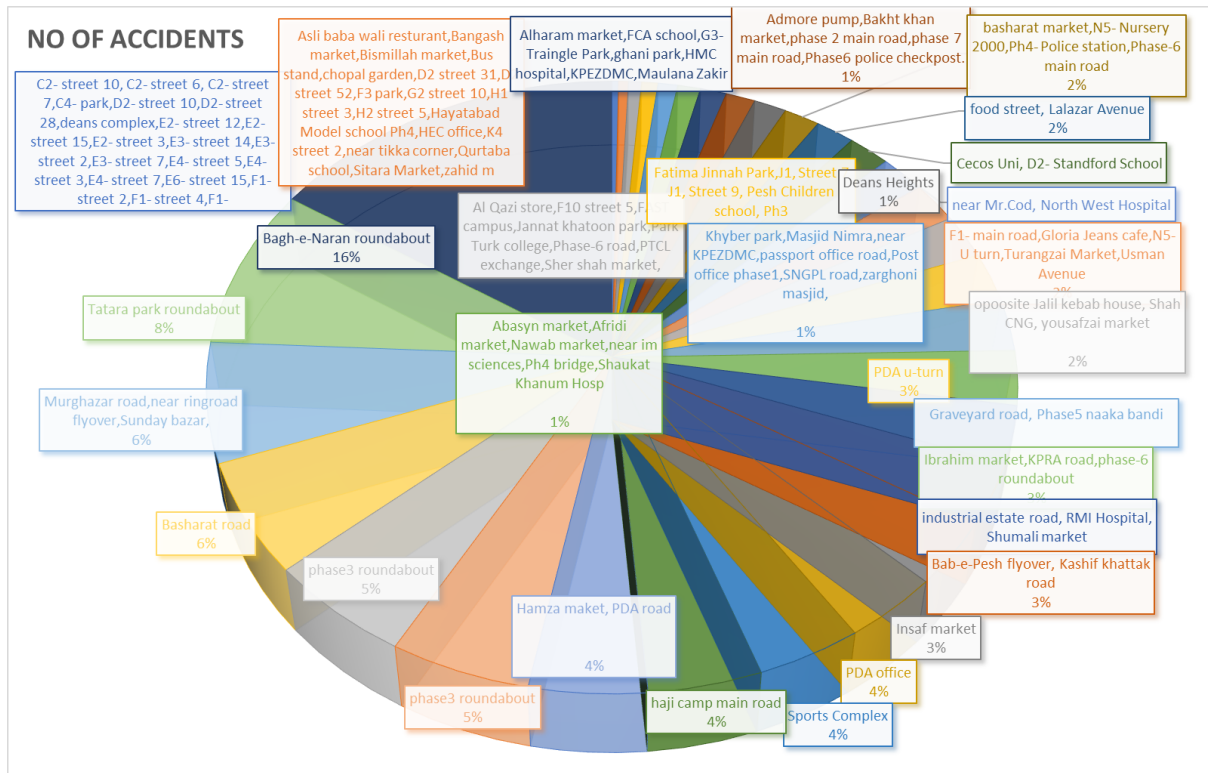
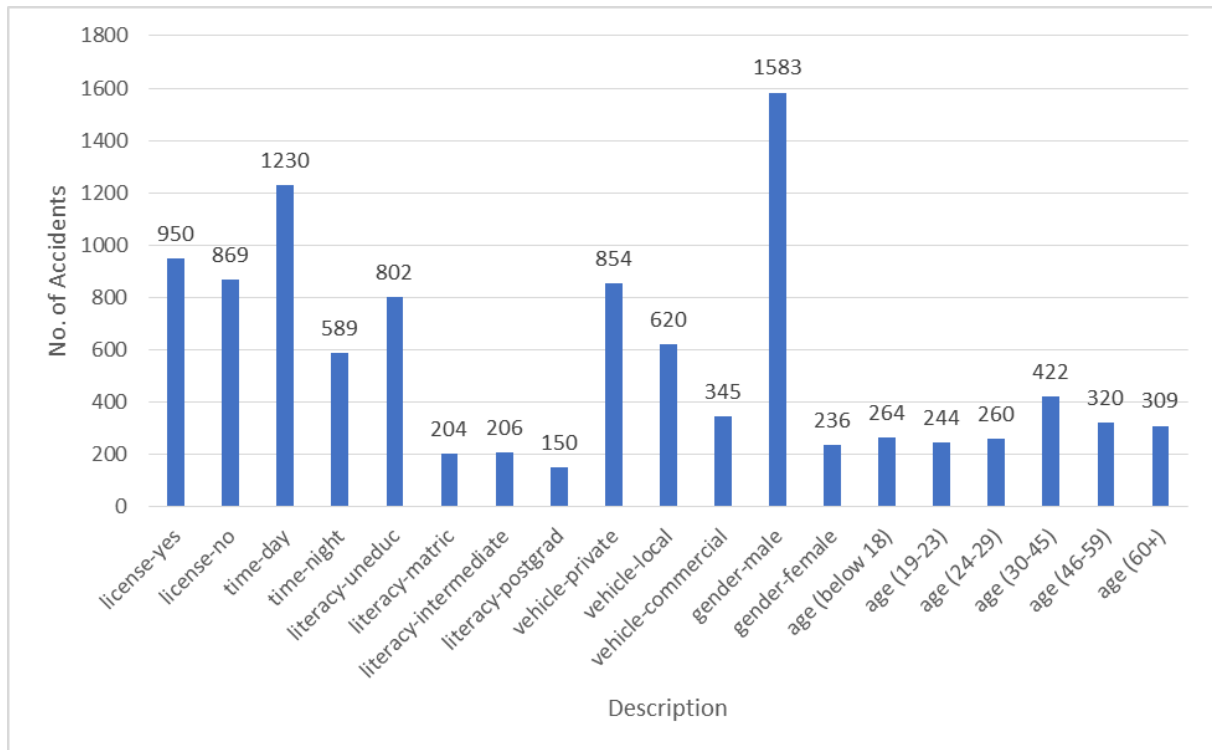


Figure 6. Accidents locations w.r.t percentage







**Figure 8.** Description vs. no. of accidents from 2009 to 2020

The cartographic representation takes the visualization to a whole new level. In most of the times, the accidents are illustrated in maps that are easy to identify the locations and anyone having zero knowledge of statistical data can identify the accident hotspots easily. Therefore, the need for reading notes is not needed anymore. The map was developed by Rabbani, et al. [50] using ArcGIS, this research has updated the map with recent RTAs data, which indicates the number of accidents along with the locations. The red dots show the accident occurring hotspots (accident with the highest frequency) based on the collected data, as shown below in Fig. 7.

To visualize each column of the historical data of RTAs of the Hayatabad area, this study has made it easier for traffic officials to know the number of accidents in each category. In other words, the categorical representation of the overall accidents along with its description conveyed the much-hidden information. It can be seen in Fig. 8 that people with driver's license were more involved than people having no driver's license. Similarly, most accidents occurred in the daytime as compared to the

nighttime with numbers 956 and 413 respectively. Moreover, uneducated drivers were more involved in accidents with 599 cases reported, while the postgraduate literacy level people were minimum among its other literacy class competitors of 82 people. Additionally, private car accidents were 778, which was the highest of local and commercial accidents combined. The male drivers were mostly involved in accidents with the number of accidents reaching up to 1249 people while 120 female drivers experienced accident cases. Finally, the age group of 30-45 years old road users' crashes were 453, which were highest among the other age groups. Although, the study highlights the facts based on the gathered data, yet a detailed investigation is further required to assess the human behaviour in causing road accidents. Behaviour plays an important role when an accident occurs. The study can further be extended on a larger scale which makes it easier to highlight the appropriate causes of the accidents. An infographics representation of all the data was also made for better understanding which is presented in Fig. 9 which makes it easier to understand.

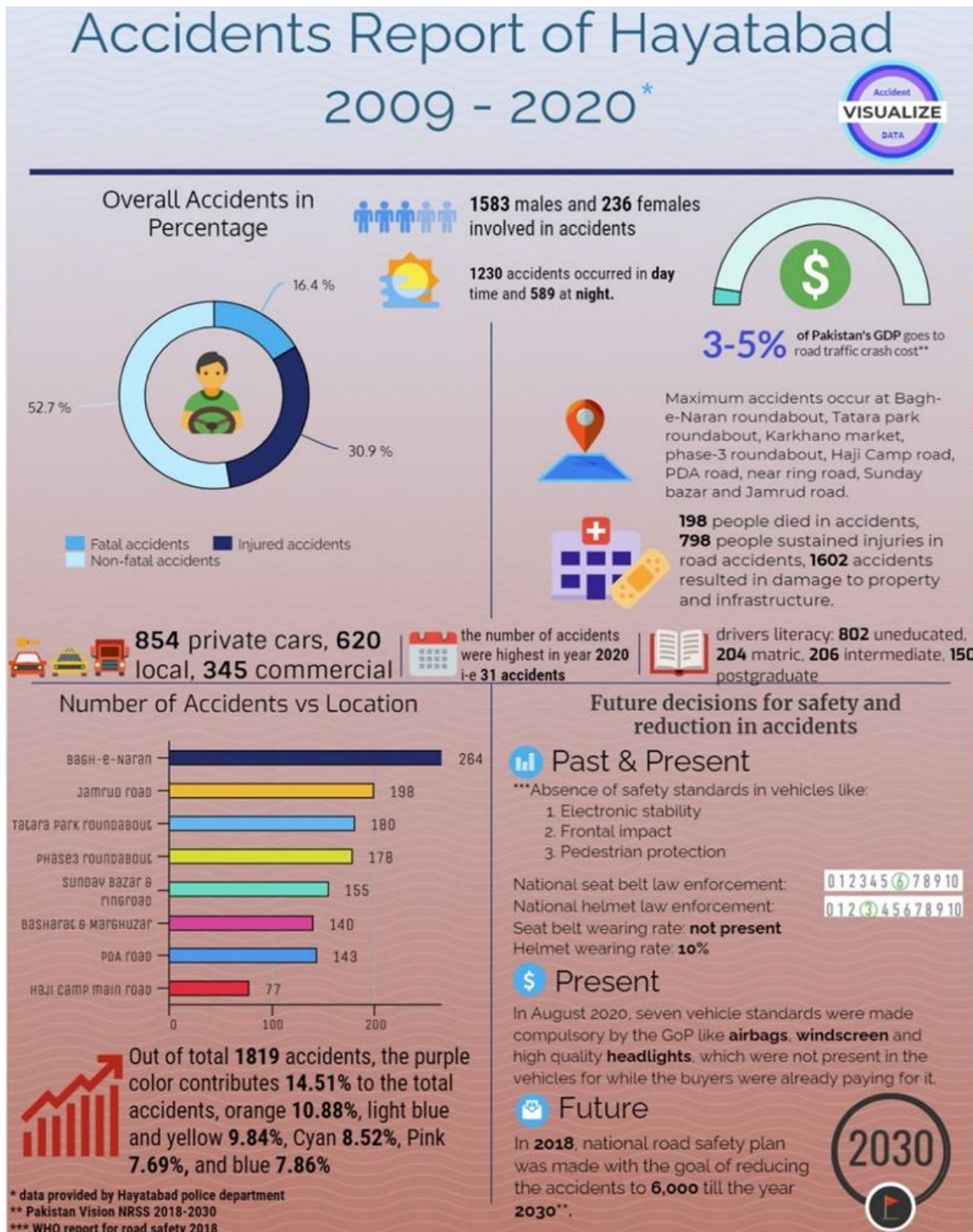


Figure 9. Infographic detail of Hayatabad road accidents

## 5. Conclusions

This study used the Road Traffic Accidents (RTAs) data for visualization. It was concluded that the RTAs analysis and the data visualization can help the planners, engineers, policymakers and government to devise

effective prevention and safety plans after studying the visualized data. In other words, it can make long-lasting and robust strategic plans to identify flaws in the current road system and traffic laws. For this purpose, the involvement at a political level, government commitment, stakeholder's allegiance and the role of social media

influencers are essential to play their role for awareness, prevention and reduction of RTAs cases as the citizens deserve better and safe road usage. The visualized data can be utilized fruitfully for transport officials and accident prevention and safety department as they can reliably extract useful information about the critical locations where most RTAs occur. By knowing the hotspots, the concerned department can implement pre-crash and post-crash safety and traffic controlling devices. Also, accident awareness and safety programs can be arranged to promote safety awareness. The main contribution of this study is to use visualization of the collected data to investigate and identify the factors contributing significantly to accidental happenings in most frequent locations and identifying the time and gender in most of the RTAs.

## Acknowledgments

The authors would like to thank Universiti Teknologi PETRONAS (UTP) for the support provided for this research.

## REFERENCES

- [1] D. Hartanti, R. N. Aziza, and P. C. Siswipraptini, "Optimization of smart traffic lights to prevent traffic congestion using fuzzy logic," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 17, no. 1, pp. 320-327, 2019.
- [2] M. B. A. Rabbani, M. A. Musarat, W. S. Alaloul, M. S. Rabbani, A. Maqsoom, S. Ayub, H. Bukhari, and M. Altaf, "A Comparison Between Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ES) Based on Time Series Model for Forecasting Road Accidents," *Arabian Journal for Science and Engineering*, pp. 1-26, 2021.
- [3] H. Bukhari, M. J. Thaheem, M. A. Musarat, W. S. Alaloul, and M. Altaf, "Are Pakistani construction professionals truly happy? A benchmarking approach," *Ain Shams Engineering Journal*, 2021.
- [4] H. A. Aziz, S. V. Ukkusuri, and S. Hasan, "Exploring the determinants of pedestrian-vehicle crash severity in New York City," *Accident Analysis & Prevention*, vol. 50, pp. 1298-1309, 2013.
- [5] D. Santoso and T. Tjahjono, "Determination minimum distance between ramp entry and exit on the freeway reviewed from the road safety aspect," *Civil Engineering and Architecture*, vol. 8, no. 3, pp. 268-273, 2020. DOI: 10.13189/cea.2020.080309.
- [6] H. Lenora Leuhery, "Determination of Black Site Area Based on Equivalent Accident Number Analysis: Case Study National Roads in Ambon City," *Civil Engineering and Architecture*, vol. 8, no. 5, pp. 1063 - 1073, 2020. DOI: 10.13189/cea.2020.080533.
- [7] W. Rafiq, M. Napiah, N. Z. Habib, M. H. Sutanto, W. S. Alaloul, M. I. Khan, M. A. Musarat, and A. M. Memon, "Modeling and design optimization of reclaimed asphalt pavement containing crude palm oil using response surface methodology," *Construction and Building Materials*, vol. 291, p. 123288, 2021.
- [8] C. Zhang, X. Yan, L. Ma, and M. An, "Crash prediction and risk evaluation based on traffic analysis zones," *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [9] L. J. Blincoe, A. G. Seay, E. Zaloshnja, T. R. Miller, E. O. Romano, S. Luchter, and R. S. Spicer, "The economic impact of motor vehicle crashes, 2000," United States. National Highway Traffic Safety Administration 2002.
- [10] W. S. Alaloul, M. A. Musarat, H. Mehmood, and M. Altaf, "ASSESSMENT OF LABOUR PRODUCTIVITY IN ROAD CONSTRUCTION PROJECTS OF PAKISTAN," *Journal of Civil Engineering, Science and Technology*, vol. 12, no. 1, pp. 32-38, 2021.
- [11] W. S. Alaloul, M. Altaf, M. A. Musarat, M. Faisal Javed, and A. Mosavi, "Systematic Review of Life Cycle Assessment and Life Cycle Cost Analysis for Pavement and a Case Study," *Sustainability*, vol. 13, no. 8, p. 4377, 2021.
- [12] W. Rafiq, M. A. Musarat, M. Altaf, M. Napiah, M. H. Sutanto, W. S. Alaloul, M. F. Javed, and A. Mosavi, "Life Cycle Cost Analysis Comparison of Hot Mix Asphalt and Reclaimed Asphalt Pavement: A Case Study," *Sustainability*, vol. 13, no. 8, p. 4411, 2021.
- [13] T. Anjuman, S. Hasanat-E-Rabbi, C. K. A. Siddiqui, and M. M. Hoque, "Road traffic accident: A leading cause of the global burden of public health injuries and fatalities," in *Proc. Int. Conf. Mech. Eng. Dhaka Bangladesh*, 2020, pp. 29-31.
- [14] M. Peden, R. Scurfield, D. Sleet, A. A. Hyder, C. Mathers, E. Jarawan, A. Hyder, D. Mohan, and E. Jarawan, *World report on road traffic injury prevention*. World Health Organization, 2004.
- [15] W. H. O. (WHO). (2018). *Global Status Report on Road Safety 2018*. Available: [https://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2018/en/](https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/)
- [16] WorldAtlas. The 10 Largest Cities in the World. [Online]. Available: <https://www.worldatlas.com/articles/the-10-largest-cities-in-the-world.html>
- [17] R. S. Margie Peden, D. M. David Sleet, E. J. Adnan A. Hyder, and a. C. Mathers. World report on road traffic injury prevention [Online].
- [18] U. d. o. Transportation. Traffic Safety Facts [Online]. Available: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812580>
- [19] A.-T. A. Jacobs G, Astrop A. , " Estimating Global Road Fatalities," Accessed on: 26-Aug-2020 Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.174.5207&rep=rep1&type=pdf>
- [20] S. R. Peden M, Sleet D, Mohan D, Hyder AA, Jarawan E, et al. World report on road traffic injury prevention. Geneva [Online]. Available: [https://www.who.int/violence\\_injury\\_prevention/publications/road\\_traffic/world\\_report/en/](https://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en/)
- [21] H. Ritchie. "Causes of Death". [Online]. Available:

- <https://ourworldindata.org/causes-of-death>
- [22] W. H. Organization. (2020, 27-Aug-2020). *Road traffic injuries*. Available: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>
- [23] [23] WHO, "The top 10 causes of death," 2020.
- [24] [24] WHO. (2020). *PAKISTAN: ROAD TRAFFIC ACCIDENTS*. Available: <https://www.worldlifeexpectancy.com/pakistan-road-traffic-accidents>
- [25] WHO. (2020). *INDIA: ROAD TRAFFIC ACCIDENTS*. Available: <https://www.worldlifeexpectancy.com/india-road-traffic-accidents>
- [26] WHO, "CHINA: ROAD TRAFFIC ACCIDENTS," 2020.
- [27] WHO. (2020). *IRAN: ROAD TRAFFIC ACCIDENTS*. Available: <https://www.worldlifeexpectancy.com/iran-road-traffic-accidents>
- [28] WHO, "TAJKISTAN: ROAD TRAFFIC ACCIDENTS," 2020.
- [29] WHO. (2020). *AFGHANISTAN: ROAD TRAFFIC ACCIDENTS*. Available: <https://www.worldlifeexpectancy.com/afghanistan-road-traffic-accidents>
- [30] W. L. Expectancy. (2020). *PAKISTAN: ROAD TRAFFIC ACCIDENTS*. Available: <https://www.worldlifeexpectancy.com/pakistan-road-traffic-accidents#:~:text=Pakistan%3A%20Road%20Traffic%20Accidents&text=According%20to%20the%20latest%20WHO,Pakistan%20%2395%20in%20the%20world>
- [31] P. B. o. Statistics. (2020). *Traffic Accidents*. Available: <http://www.pbs.gov.pk/sites/default/files/tables/renam-as-per-table-type/Traffic%20accidents%20%28Yearly%29.pdf>
- [32] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE access*, vol. 2, pp. 652-687, 2014.
- [33] M. Razak, S. Fauzi, R. Gining, and M. Jamaluddin, "Data visualisation of vehicle crash using interactive map and data dashboard," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 3, pp. 1405-1411, 2019.
- [34] O. Kotavaara, H. Antikainen, and J. Rusanen, "Urbanization and transportation in Finland, 1880–1970," *Journal of Interdisciplinary History*, vol. 42, no. 1, pp. 89-109, 2011.
- [35] Wikipedia. (2020, 29-Aug-2020). *Python (programming language)*. Available: [https://en.wikipedia.org/wiki/Python\\_\(programming\\_language\)#cite\\_note-About-29](https://en.wikipedia.org/wiki/Python_(programming_language)#cite_note-About-29)
- [36] Python. (2020, 29-Aug-2020). *About Python*. Available: <https://www.python.org/about>
- [37] E. Debill. (2019, 29-Aug-2020). *ModuleCounts*. Available: <http://www.modulecounts.com>
- [38] A. Rao. (2020, 29-Aug-2020). *Top 10 Python Libraries You Must Know In 2020*. Available: <https://www.edureka.co/blog/python-libraries/>
- [39] A. Bronshtein. (2017, 29-Aug-2020). *A Quick Introduction to the "Pandas" Python Library*. Available: <https://towardsdatascience.com/a-quick-introduction-to-the-pandas-python-library-f1b678f34673>
- [40] G. R. McKenzie, "Data charts: A crutch for helping pupils organize reports," *Language Arts*, vol. 56, no. 7, pp. 784-788, 1979.
- [41] M. Patel, A. Khalifa, M. Liew, Z. Mustaffa, and A. Whyte, "Downtime cost analysis of offloading operations due to influence of partially standing waves in Malaysian waters and development of graphical user interface," *Annals of Operations Research*, pp. 1-27, 2020.
- [42] P. Ljubič, L. Todorovski, N. Lavrač, and J. C. Bullas, "Time-series analysis of uk traffic accident data," in *Proceedings of the Fifth International Multi-conference Information Society*, 2002, pp. 131-134.
- [43] U. Lokala, S. Nowduri, and P. K. Sharma, "Road Accidents Bigdata Mining and Visualization using Support Vector Machines," 2017.
- [44] K. J. Renze, "Accident/Incident Reconstruction and Visualization," in *ASME/IEEE Joint Rail Conference*, 2018, vol. 50978, p. V001T02A015: American Society of Mechanical Engineers.
- [45] R. Poudel, "Network Traffic Visualization with Scapy and ELK Stack," undergraduate, Computer Science, London Metropolitan University, 2019.
- [46] W. Chen, F. Guo, and F.-Y. Wang, "A survey of traffic data visualization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2970-2984, 2015.
- [47] C. Silva, "Urban Traffic Visualization at UT Austin Portugal Applied Visualization," ed, 2019.
- [48] A. Elhassan, "Epidemiology of Road Traffic Accidents in Sudan," *European Academic Research*, 2020.
- [49] J. Sodikov, "Road Traffic Accident Data Analysis and Visualization in R," *International Journal of Computer Science Engineering and Information Technology Research (IJCEITR)*, 2018.
- [50] M. B. A. Rabbani, S. A. Khan, Q. Iqbal, and E. Q. Zaman, "Analysis of Vehicle Accidents using Spatio-Temporal Tools in ArcGIS: A Case Study of Hayatabad, Peshawar," 2019.