Visual Data Analysis

UNIVERSITY
OF SKÖVDE

Analysis of Road Accidents in India

Divya Chedayan

Student ID : a22divch

Visual Data Analysis report
Spring term 2023

Table of Contents

1. Description and Background

India is one among the most populated countries in the world. Currently, India has the second largest population in the world, with 1,419.66 million people [5]. India having one of the largest road network, road accidents have become a significant concern due to many reasons. According to the Ministry of Road Transport and Highways, India saw 449,002 road accidents in 2019, with 151,113 deaths and 451,361 injuries. There is an increase in number of road accidents happening in India from 1970 and due to this huge increase in the road accident rates, huge amounts of data are generated keeping records of these accidents. The causes of road accidents could be different factors which include the feature of the road, weather conditions, fault of drivers, not following traffic rules, not following safety measures, and so on. In this project, I would like to study and analyze the accidents that happened in India in both national and states level during the period between 1970 and 2019 to understand the trend and the significant causes of road accidents throughout the period. As per the data, there has been a massive increase in the number of accidents in India from 1970 to 2019.

The increase in the population and motorization in the country, along with the expansion of the road network, led to increased road accidents all over the country, getting many injured and even causing deaths. Visualizing each dimension of the road accident could help the authorities to create awareness and decision-making to help reduce the number of accidents. Here, I am analyzing the data provided by the Indian Government between a fixed period from 1970 to 2019 to find out the trends of accidents happening during this period, and I will be more concentrating towards the latest data provided from 2016 to 2019 to understand the primary cause of these accidents and what other factors could be taken into account to decrease these accidents in India nationally, in different states and union territories (unlike states, union territories are directly ruled by the Central Government, and there are 8 of them), and cities. Visualizing a large set of data helps to understand various factors and find out the relationship between these factors and the contribution of these factors leading to the severity of accidents. Visualizing the data helped to understand that overspeeding is the primary cause of road accidents leaving many injured and killing many lives.

WHO [7], in their Global status report on road safety in 2018, the annual count of road accidents in India reached 1.35 million. According to them, injuries due to road accidents are the leading killer of people aged between 5 – 29 years. After analyzing the data, numerous factors contribute to the high rate of road accidents in India, including poor road infrastructure, poor traffic management, lack of enforcement of traffic rules, heavy loaded vehicles, driver's fault, and poor vehicle maintenance. Another significant factor is the high number of two-wheelers on Indian roads, which are more vulnerable to accidents compared to other vehicles. The absence of safety gadgets such as helmets can lead to death. Moreover, the fault of the

drivers are a leading cause of road accidents. There are numerous drivers influenced by alcohol or drugs or while fatigued could lead to worst driving leading to road accidents. Overall, there is a need to improve the road safety in India including better infrastructure, stricter traffic laws, and creating awareness among road users about the importance of safe driving.

1.1 Summary of two papers

In this section, I will provide a brief description of the research papers that provided information regarding road accidents in India.

According to Dinesh Mohan [3], during the last past ten years, road fatalities caused due to accidents have increased by 8% every year, and not been any decrease in the figure each year. Dinesh Mohan, in his advanced research on Road safety in India [2], examined the situation of traffic safety in India and analyzed the areas that caused severe accidents causing major harm, and suggested ways to decrease the accidents. They mainly focused on the two aspects of traffic safety, i.e, challenges and opportunities. Firstly, the report provided the current traffic situation in India, and they found there is a major increase in fatality rates on both the highways and the urban areas. Secondly, they provide a few measures that will help to decrease the fatalities caused due to road accidents. Finally, in their report, the researchers compared the important traffic challenges between India and China.

Sanjay Kumar Singh (2016) [6] also studied the issues and challenges of road traffic accidents in India. In his research, he analyzed road accident deaths and injuries based on people's age, gender, month, and time. He found out there are higher male fatalities compared to women, and also road accidents are seen more in May – June and December – January due to weather conditions. He also analyzed the trend in the accident in 2013, and the different aspects that led to cause the accidents.

Dr Jayaprakash G Hugar [4] study was conducted on the scientometric analysis of the road accidents happening in India from the period 1970 to 2020. The researchers defined accidents as events that happen in a sequence that mostly lead to unintended injury, death, or damage to property. According to them, today's accidents are the most dangerous and leading cause of death all around the world.

1.2 Motivation of your visualization

Many research were performed to reduce road accidents in India and they have suggested solutions to reduce them. Still today there is not much improvement seen in the road infrastructure and more accidents are happening every year that

is leading to many fatalities as well. When analyzed the data, I found there is no decrease in the number of fatalities each year, so wanted to find out the real causes of these accidents. Analyzing the trend of road accidents happening in India to suggest measures to reduce them. Visualization helps to analyze the data of road accidents over the period of time. By comparing between the number of accidents and related deaths in 2019 with the data of previous years, it is easier to identify whether the situation is improving or worsening, thereby helping to understand the underlying causes of road accidents and take steps to reduce the number of accidents in future.

Most research was mostly finding the causes leading to road accidents. I couldn't find studies on urban and rural areas combined with medical care. I found this interesting to explore more data to provide solutions for improving medical facilities in both urban and rural sectors. Visualization can be a powerful tool to help people understand the scope and the severity of a problem. Visualizing road accidents and their related injuries and deaths helps to create awareness and encourage people to take them seriously. Visualization leads to identifying the areas with the highest number of road accidents that can help the authorities to take safety measures and invest in the road infrastructure to reduce them. In the referred research papers, they used bar charts, line charts, and pie charts to visualize the data until 2018. So, I would like to include more types of charts for better multivariate visualization, which could be done with the help of Spotfire focusing more on 2019 data.

2 Data

The dataset is downloaded from https://data.gov.in/catalog/road-accidents-india-2019. The information is provided by Indian Government (Ministry of Road Transportation) on their official website, which include all the details regarding road traffic accidents from the period 1970 to 2019. The data includes all the details of accidents happened in India as well as in each states and Union Territories. There are total count of accidents that happened, a count of injured and deaths caused due to road accidents, weather conditions, people's gender, and age, types of vehicles increasing the severity of accidents, etc.

3 Research questions (or hypothesis)

- Trend of accidents happening all over India from 1970 to 2019 and also comparing different factors leading to road accidents nationally and in each states and union territories.
- What are the major causes of road accidents in India?
- Is weather an important factor for road accidents?
- Which type of vehicles encounter more accidents and cause fatalities?
- Does our health care system helps to decrease the fatalities in urban and rural areas of India?

- Age and Gender of people involved in road accidents and which gender is most affected?

4 Visualization Results

4.1 Visualization

Visualization helps us to visually represent a huge amount of data to view them in a better way. It helps us to understand the data better and study the impact on the project which helps in proper decision-making.

4.1.1 Analyzing historical data

RQ1 : Find the trend or pattern of road accident occurrences in India.
In my project, I have started with analyzing the data from 1970 to 2019 for a better understanding of the trend of accidents that happened in India. I have chosen bar chart to show the trends in a simple way which includes various features for better understanding. As we might know and based on the data, the registered vehicles were significantly fewer in the 90s compared to today, and due to that, the count of accidents that occurred was low.

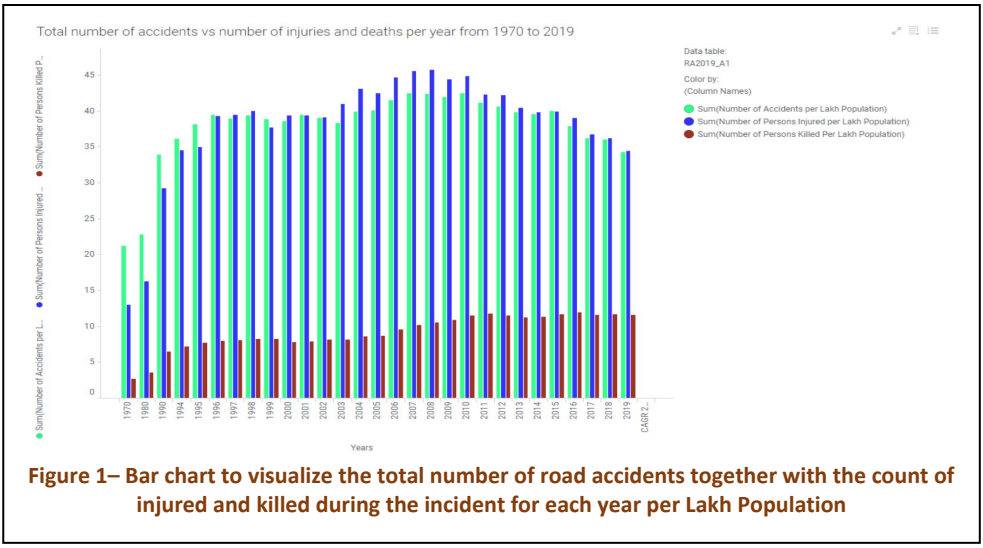


Figure 1– Bar chart to visualize the total number of road accidents together with the count of injured and killed during the incident for each year per Lakh Population

Observation : The bar graph clearly represents there is an increment in road accidents until 2010, and after that there is a slight decrease in the number of accidents until 2019 and the count of injured people are high with minor or greviously injured. But overall the number of deaths remains low, but the graph showing increment each year. This could be due to large increase in registered vehicles throughout the period. During this period the registered vehicle count has increased from 1401 to 297200 by 2019.

4.1.2 Causes of Road Accidents

RQ2: What are the major causes of road accidents in India?

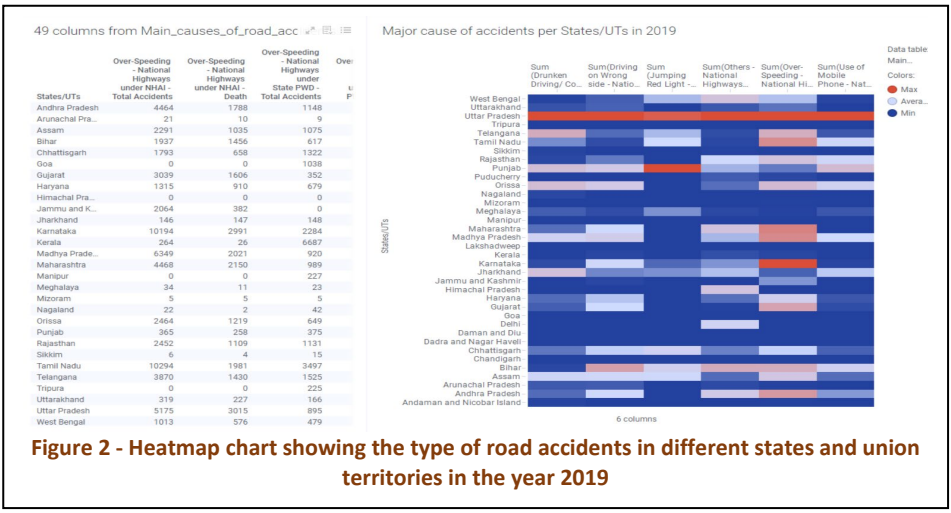


Figure 2 - Heatmap chart showing the type of road accidents in different states and union territories in the year 2019

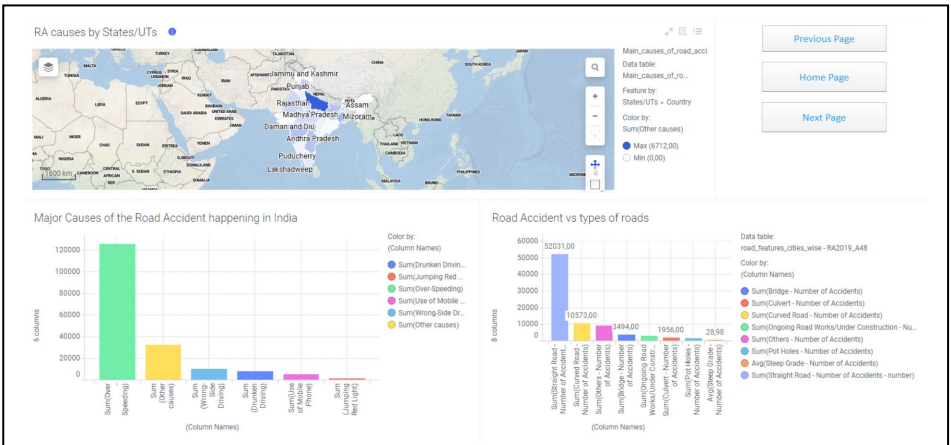


Figure 3 – Bar charts to visualize the major causes of road accidents in India and map chart showing states/UTs wise having the highest number of different causes. This map shows the states/UTs with maximum number of accidents due to wrong side driving.

Observation : Overspeeding is the major cause of road accidents in India. Also, more than half of the accidents occurred in a straight road, which clearly shows the correlation with overspeed of the vehicle. The most appropriate solution to reduce the speed could be setting up a speed limit and surveillance cameras to watch vehicles passing in straight roads. Applying speed-breaker could also help to reduce them.

The above graphs shows what could be the main causes of road accidents in different states and union territories of India. Heatmap gives a better visualization which

clearly shows the states/UTs with highest number of road accidents based on different causes. It is clear that the primary cause is the overspeed of the vehicle. Other reasons include driving after intake of alcohol or drugs, driving on wrong side of the road, jumping red light, using of mobile phones, driver distraction, and so on. I have chosen bar charts to show the data in simple way and identify the major causes that leads to road accidents. Uttar Pradesh has the highest number of accidents in 2019. Followed by Uttar Pradesh, Karnataka has highest number of accidents due to overspeeding. Jumping red light at the traffic signals is the major cause of road accidents in Punjab. States like Tamil Nadu, Bihar, Orissa, Madhya Pradesh has more accidents causing injuries and deaths due to these factors.

4.1.3 Effects of Weather and Seasons on Road Accidents

RQ3 : Is weather an important factor for road accidents?

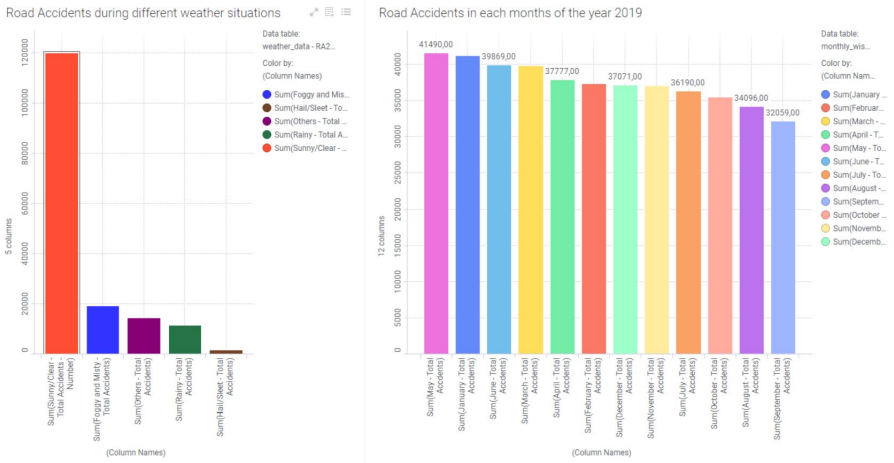


Figure 4 – Bar charts showing the number of road accidents happened in different weather condititions and comparing the accident cases in each month (January – December) in 2019

I have used bar charts to show the accidents happened in different weather conditions and which months have high number of accidents, which is a simple and clear way to visualize the weather data and compare between each month in 2019.

Observation : The Figure 4 shows 60 % of the accidents occurred when the sky is sunny or clear. Only fewer accidents due to weather conditions such as fog/mist, rainy and hail/sleet. This clearly explains that the weather conditions are not a major factor leading to accidents. India is a large country with large number of population and each different parts of India have different types of weathers. Nothern regions are more colder, and winter months are covered with snow and mostly

foggy throughout the year. More rain could be seen in the southern part. The second bar chart explains the number of accident cases happened in each month from January to December in the year 2019. January and May are two critical months with maximum number of accidents faced by the states in India. The most number of accidents are in winter and summer, two main seasons of India. I would like to conclude that weather is not a critical factor in causing accidents in India, but extreme weather condition might cause more accidents according to the data. May be drivers are more cautious during driving in bad weather, while accidents occurred in clear or sunny weather may be due to overspeeding and violations of traffic rules.

4.1.4 Road Accidents with Types of Vehicles

RQ4 : Which type of vehicles encounter more accidents and cause fatalities?

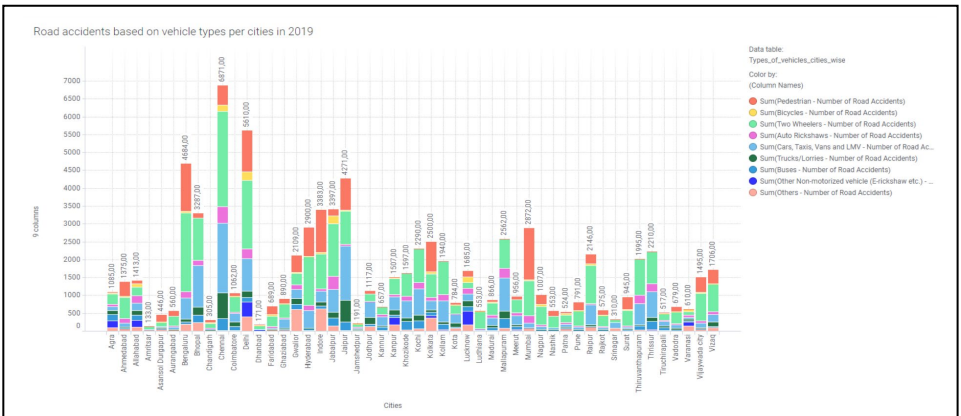


Figure 5 - Stacked bar chart to compare the number of accidents based on different types of vehicles

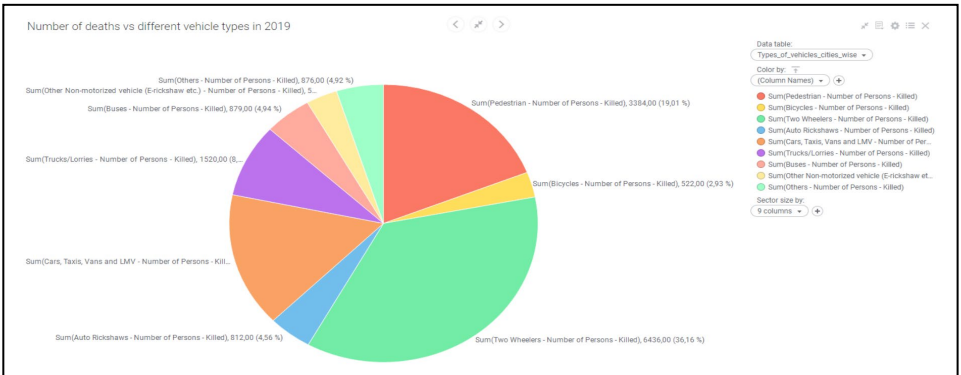


Figure 6 - Pie chart displaying the number of fatalities with different types of vehicles in 2019

I have used stacked bar chart which helps to compare different vehicle types for each states/UTs to identify which vehicle type is not safe while travelling and thus generating awareness of following safety rules and gadgets. The pie chart helps to show the proportion of data that clearly explains the relationship between the vehicle types and road accidents.

Observation : The stacked barchart shows the types of vehicles that encountered accidents in each metropolitian cities of India. This graph says that the people travelling in vehicles such as two-wheelers, cars, taxis and vans and pedestrians are the others suffering the most. Approximately 19 % of pedestrians were killed in the road accident in 2019. This could be because of any road construction works or not enough facilities of footpaths on the sides of the roads, and lack of road infrastructure and traffic signals. According to the data, the majority of people around 36.16 % killed using Two-wheelers. Cars, taxis and vans are also other main types of vehicles facing accidents.

4.1.5 Healthcare and road accidents

RQ5 : Does our health care system helps to decrease the fatalities in urban and rural areas of India?

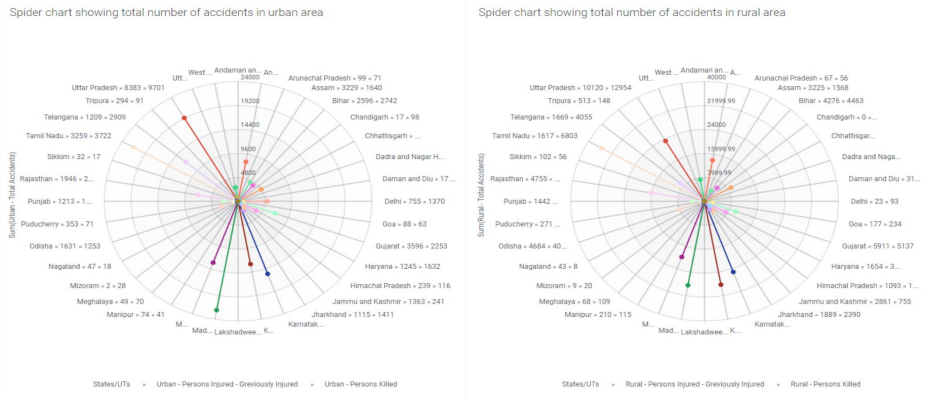


Figure 7 – Spider chart showing the details of the accident with people injured grievously and killed in both urban and rural areas

I have used radar or spider chart to visualize multivariate features which includes the different aspects of urban and rural areas. This graph include the states/UTs, the count of people injured greviously and the total number of deaths to measure the efficiency of the medical care available in India. The previous research papers compared different bar charts to analyze this data. I found visualizing in spider chart more efficient to analyse each states/UTs with the measurement of other necessary fields together. Also, comparing the cases in urban and rural areas is easy using

spider chart. A scrollable bar chart is used to compare the fatalities among urban and rural areas.

Observation : According to the data provided by the Indian Government, majority of states have decrease count of the fatalities caused by the road accident. It is clear that there are more accidents occurring in urban area than the rural areas in most states/UTs and a few states/UTs have same count of accidents in both urban and rural areas. Even though states have a high number of accidents causing injuries to people greviously, the death rate is very low. This means the heath care system is strong in India, mostly in the urban areas, and this could be extended more towards the rural areas for better treatment of the rural people.

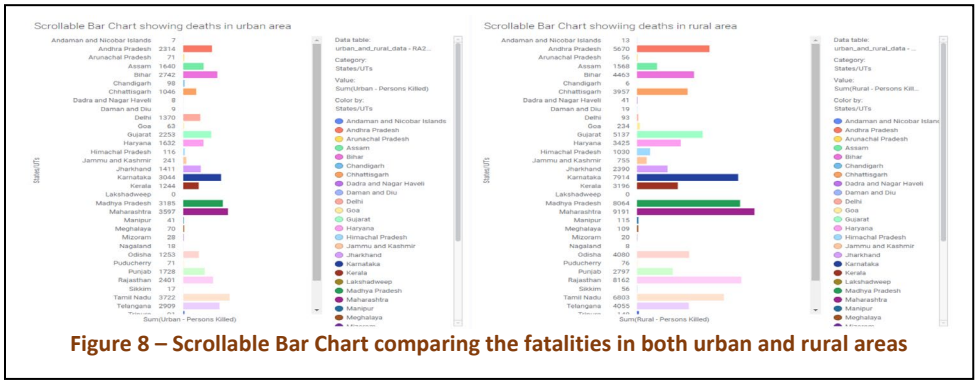


Figure 8 – Scrollable Bar Chart comparing the fatalities in both urban and rural areas

An increase in development of hospitals and emergency clinics in rural areas can again decrease the fatalities count. I couldn't find the exact data on the healthcare system related to road accidents. Still, the features used to visualize the above plot clearly provide information regarding the person's injury (minor or major) and fatalities.

4.1.6 The age and gender of the people involved in road accidents

RQ6 - Age and Gender of people involved in road accidents and which gender is mostly affected?

Visualizing in a pie chart below allow us to show the proportion of each age group. And its easy to compare among each gender type and age group.

Observation : The data shows that nearly 80 % of males are involved in road accidents, which includes all age groups. The number of females encountering road accidents is very low compared to males in 2019.

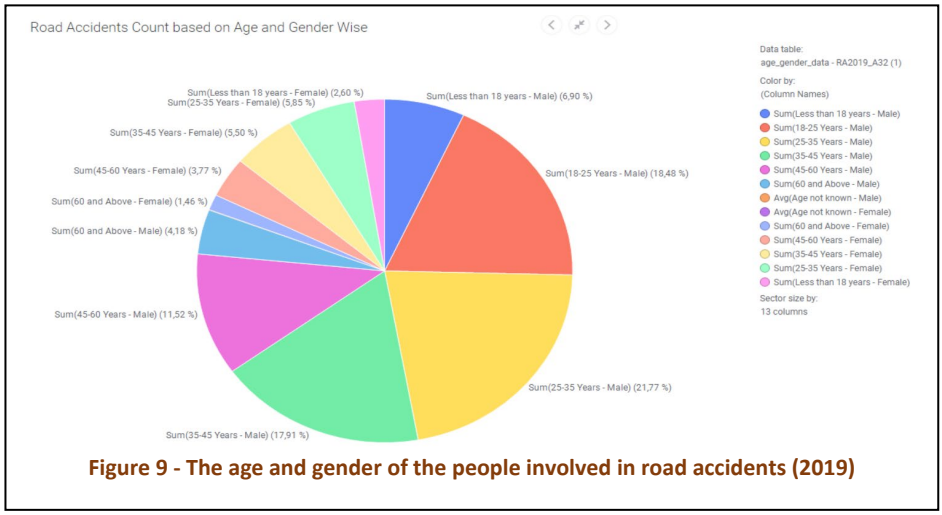


Figure 9 - The age and gender of the people involved in road accidents (2019)

4.2 Data Analysis (What, Why, How)

The approach is to find the answers to the mentioned hypothesis in Section 3.

What : The data collected from Ministry of Road Transportation and Highway website by Indian Government, which provide huge data related to road accidents in India. There were abundance of data related to the road accident occurred from 1970 to 2019. I took the data from 1970 to learn the trend of the road accidents happened, that helped to visualize the lastest data mainly focusing on 2019 data. Selecting the appropriate dataset from the whole data was time consuming one. Prioritizing the data and removing the data which is not useful to answer different problem statements were performed. Also, data were splitted into different datasets which needs to be combined in order to solve some of the problems. I have also noticed there are missing values in the dataset, thus cleaning of data is the most important thing before visualizing the data.

Why : After analyzing the data, it shows there was large increase in the number of accidents from 1970 to 2010. After 2010, the data shows a slight decrease in the count of accidents until 2019. But the number of deaths is increasing every year. To reduce the number of accidents and deaths, various factors were analyzed to understand the significant causes of road accidents, thereby helping to take measures to save the lives of people.

How : The visualizing of data helps for the better understanding of data and finding insights that will help to measure the severity of accidents, which indeed helps to take measures to prevent road accidents in India. Once data is prepared, we can

visualize the data. Here, I have used a bar chart, stacked bar chart, Heatmap, pie chart, spider chart, scrollable bar chart, and map.

The historical data is visualized using a bar chart with years in the x-axis and the total number of accidents, and the number of people injured and killed during accidents in each year. Although there is small visible decrease in the number of accidents between 2010 to 2019, the number of fatalities is increasing every year. To find the causes of the accidents at the state level, a parallel coordinate chart and bar charts are used to visualize, helping to answer the leading causes. The parallel coordinate chart is used to show the effect of different causes of road accidents in every state. The bar charts are used to find the primary cause of road accidents and how road features are a major factor causing road accidents.

To analyze how the weather and seasons can cause road accidents, bar charts are used. Also visualized the data month-wise to see how the weather and season can cause road accidents. After analysis, weather could not be a reason for road accidents because the majority of accidents took place in the clear and sunny sky. Analyzing month-wise data, January and May show the highest number of accidents. Overall, extreme weather conditions could be a factor in leading to road accidents.

With the help of a stacked bar chart and pie chart, I was able to visualize the road accident based on types of vehicles. The stacked bar chart clearly shows the distribution of each vehicle types in different states, and understood that two-wheelers are the one with highest risk of getting into an accident. More strict law enforment to follow safety rules could help to reduce the fatalities.

To understand the road accidents in urban and rural areas, the spider chart was used. This spider chart include the states/UTs, greviously injured count and the fatalities count. I have also used scrollable bar chart to compare the deaths in urban and rural areas which shows more deaths are in rural areas. Better medical care could be implemented in rural areas to reduce the number of fatalities. Also, there are higher male fatalities compared to females.

4.3 Reflection

One of the leading causes of injuries and deaths in India is road accidents. This is a common problem for all developing countries without a good road infrastructure. There have been many studies conducted to understand the main causes of road accidents. Based on the research, measures are taken to minimize the accidents happening. They have visualized the data using bar charts, line charts and maps and also found some have performed visualization using python visualization tools in Jupyter notebook. The visualization project attempts to understand the main causes of road accidents in India. Based on the data, overspeeding is the primary

cause of road accidents leaving many injured and deaths. The causes differ from each state. The types of roads or road features also contribute to the road accidents. Majority of accidents occurred in the straight road, and maybe the driver's fault contributed to the majority of the accidents. Uttar Pradesh had the highest number of accidents in 2019, followed by Punjab, Tamil Nadu, Karnataka, and many other states had severe accidents with high fatalities.

The study can be extended to the risk of accidents in types of junctions and take the measures to reduce the accidents, and also the risk of accidents with heavily loaded vehicles could be studied. The study of recent data could help us to understand the current pattern of accidents occurring in India. This visualization project helped me to understand the status of road accidents until 2019 with the data provided and the causes affecting them. Did the measures applied earlier helped to decrease the number of accidents? I don't think it has contributed much, but more steps should be taken to save people's life.

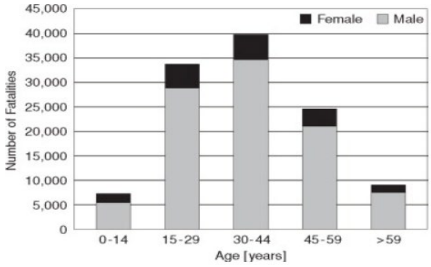
The main limitation is the availability of recent data, and the features needed to analyze the most latest trend. The data available in the Government website is until 2019. Many hotspots areas, which had highest number of accidents in each states/UTs has been identified, that can help the authorities to take safety measures. Several suggestions are mentioned in this report to reduce road accidents. Implementing these solutions will take years, such as building a good road infrastructure to implement as a developing country. Government has road safety rules which should be updated every time to lower the number of accidents. Strict traffic rules should also be implemented.

5 References

1. Chakradhara Panda, A. K. (2022, August 01). *Assessment of Risk Factors of Road Traffic Accidents: A Panel Model Analysis of Several States in India.* Retrieved from Vision: The Journal of Business Perspective: https://journals-sagepub-com.libraryproxy.his.se/doi/10.1177%2F09722629221113251
2. Dinesh Mohan, O. T. (2009, January). *Road safety in India: Challenges and opportunities.* Retrieved from UNIVERSITY OF MICHIGAN LIBRARY: https://deepblue.lib.umich.edu/handle/2027.42/61504
3. Dinesh, M. (2014, December 23). *ROAD ACCIDENTS IN INDIA.* Retrieved from Science Direct: https://www-sciencedirect-com.libraryproxy.his.se/science/article/pii/S0386111214602399?via%3Dihub
4. Dr. Jayaprakash G. Hugar, D. M. (2020, October 01). *Road Traffic Accident Research in India: A Scientometric Study from 1977 to 2020.* Retrieved from University of Nebraska Lincoln: https://digitalcommons.unl.edu/libphilprac/4325/
5. O'Neill, A. (2022, November 17). *India: Estimated total population from 2017 to 2027.* Retrieved from Statista: https://www.statista.com/statistics/263766/total-population-of-india/
6. Singh, S. K. (2017, June 08). *Road Traffic Accidents in India: Issues and Challenges.* Retrieved from ScienceDirect: https://www-sciencedirect-com.libraryproxy.his.se/science/article/pii/S2352146517307913?via%3Dihub
7. WHO. (2018, June 17). *Global status report on road safety 2018.* Retrieved from Asia Specific Observatory On Health Systems and Policies: https://apo.who.int/publications/i/item/9789241565684

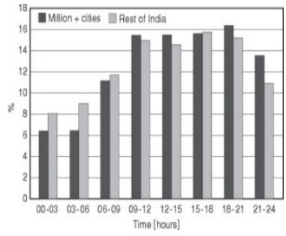
6 Appendix

Different visualization from referred research papers.



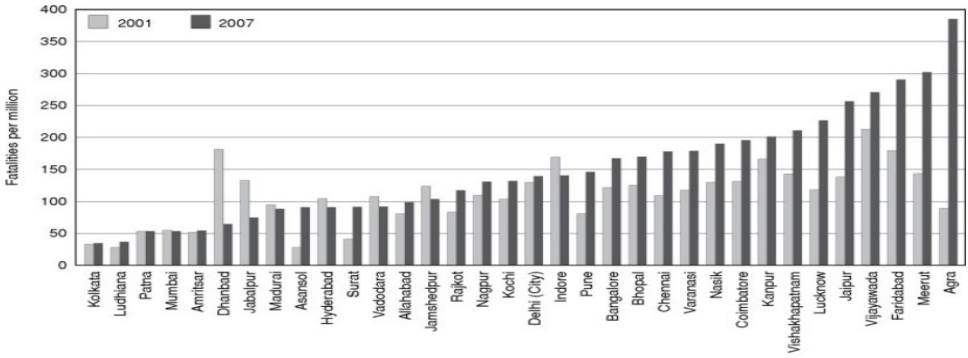
Download : Download full-size image

Fig. 1. Traffic fatalities by age and gender, India 2007



Download : Download full-size image

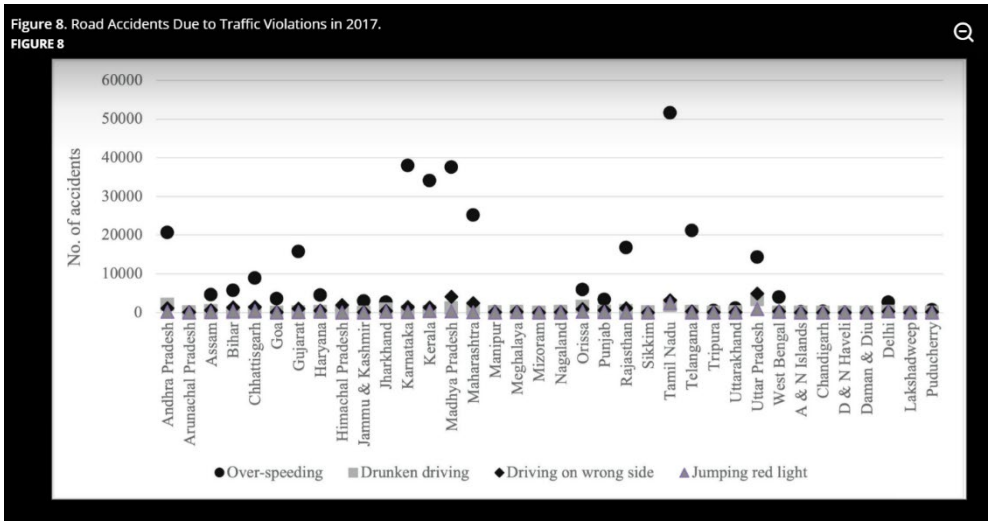
Fig. 2. Road traffic accident proportions (%) by time of day in 35 cities with more than 1 million population and those in the rest of India in 2007

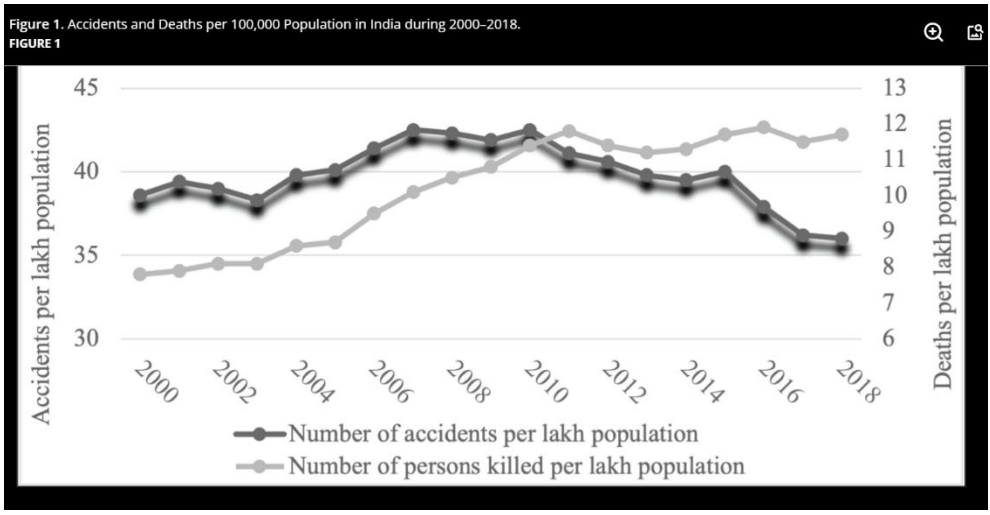


Download : Download full-size image

Fig. 3. Traffic fatality rates in cities with populations of at least one million, 2001 and 2007

Visualization graphs from Dinesh, M. (2014, December 23)





Visualization graphs from Chakradhara Panda, A. K. (2022, August 01)