

DATA ANALYTICS ON ROAD ACCIDENTS: A CRITICAL STATISTICAL REVIEW ON SAFETY MEASURES AND AWARENESS ON ACCIDENTS

S.Bharathiraja¹, Dr.P.Vaishnavi²

¹Department of Computer Applications, Anna University BIT-campus, Trichy-India,
bharathirajas1995@gmail.com

²Assistant professor, Department of Computer Applications, Anna University BIT-campus, Trichy- India,
,vaishmk@gmail.com

Abstract - Over the year, a review on safety measures and road traffic accident situation in India. This focuses on four aspects population status, road accident statistics, reason for road accidents, awareness on road accidents. The first part of report provides a comprehensive analysis of vehicle population in India. The second part of the report is pointed out in analysis that number of road accidents in the past few years. These statistics indicate the number of fatalities and number of person injured (Grievous injury/minor injury) and vehicle for road accidents in India. These are declined to continue for many years unless the new policies are implemented. The third part of the report gives a brief insights various reason for the road accidents in India. Finally the project is useful for safety measures and awareness to the people.

Keywords: Data Analytics, RStudio, RTA, Safety Measures.

I. INTRODUCTION

There are a lot of vehicles driving on the roadway every day in India, and traffic accidents could happen at any time anywhere. Some accident involves fatality, means people die in that accident. There are many phases included in, human population, vehicle population(motorcycle, car, lorry)and if it is calculated for total number of accident I the year, recent for the accidents, it provides a comprehensive analysis of report is find out the number of person killed/year and injured for (Grievous/minor) in a road traffic accidents. This project is attempt to analyze the road accident data from the year 1990 to 2016. Table1 represents the total vehicle population in India. Table2Classify the number of accident in the year. And hence finally generate the output should have at least by looking into these statistics, awareness about road safety should be given to common people. This can be done by organizing safety awareness programs across the India.

IDE-RStudio :R is an open-source software environment for statistical computing and graphics. [1] R compiles and runs on Windows, Mac OS X, and numerous UNIX platforms (such as Linux).R is designed around its core scripting language but also allows integration with compiled code written in C, C++, FORTRAN, Java, etc.A console for issuing commands. A rich set of keyboard shortcuts. Automatic source-code formatting, assistance with parentheses, keyword highlighting. RStudio provides many convenient and easy-to-use administrative tools for managing packages, the workspace, files, and more. User contributed add-on packages to supplement the base distribution. As of 2011, there were more than 3,000 such packages hosted on CRAN.

II. LITERATURE SURVEY

Assam one of the federal states in North East India has a high rate of accidents and deaths in relation to number of vehicles on the road. Here extracted detailed analysis of road traffic accidents[2]. If the complete data for RTA of the year 2009 from case dairies and police records were studied. Road traffic accidents (RTAs) have turned out to be a huge global public health and development problem killing almost 1.2 million people a year and injuring or disabling between 20-50 million people more. If it is analysis for the (rush & negligence) make 95.38%of Total RTAs.60% of the accidents were recorded during daytime (6AM to 6PM).

The peak time was between 12 PM to 6 PM (38.46%). The highest numbers of accidents (32.30%) were observed in the heavy rainy season during the months of July – September [3]. Investigation of driver behavior near accident black spots in Spain [4]. The big data analysis survey report using geo telematic data. Hundreds of thousands of people across the world lose their lives in car accidents and road disasters every year [3]. In Spain, 177 million euro and about 2 million lost working days are estimated cost of yearly traffic accidents. The figure of 177 million doesn't take into account insurance and compensation, medical costs and health resources, legal defenses, surveys, or decreased productivity in a company, among others. The major causes of accidents in human failures 80 to 90 percentage of distractions, speeding, alcohol. failures to use a seat belt maneuvers such as, harsh cornering, harsh acceleration, harsh breaking. The black spots are analyzed using number of victims and number of predominant accident type. This project shows telematics to improve upon safety and the quantity of data that is possible to process with the right tools in big data analysis.

A brief studied at all suggested that the analysis of traffic safety situation in India. [3] It could be identifies the areas in which the total harm caused by crashes can be substantially and readily reduced. In a country where 34 birth sharpening every minute and over 10 deaths happening every minute, people really don't care to mourn for the dead. Of the various reasons for a person to die, road accidents are the major ones. It is pointed out in analysis that fatality rate have increased during the past few years [5]. The analysis of road accident data 2015 reveals that about 1,374 accidents and 400 deaths take place every day on Indian roads which further translates into 57 accidents and loss of 17 lives on an average every hour in our country [6]. These statistics indicate that number of fatalities in India is not likely to start to decline for many years to come unless new policies are implemented. Second part of the paper gives a brief insight on the various reasons for road accidents in India. This project analysis of percentage share in total number of accidents [5].

According to the government of India, ministry of road transport & highways transport research wing. [4] During the year 2015, overloaded vehicles caused 77,116 accidents and 25,199 road accidental deaths. It constituted a share of 15.4 per cent and 17.2 percent respectively in total road accidents and fatalities in the country. If it is analysis for percentage of share in total number accident, number of person injured, number of killed, accident severity, vehicle for accident, causes of accidents, classification age person killed (male/female) etc. Moreover the data are represented as a percent of road traffic accident analysis and generate reports for road traffic accidents. This project initiative by the road safety and awareness gives to the common people.

A complete research studies of Road traffic injuries are the leading cause of death under 30 years, causing over a million deaths every year [5]. A road traffic injuries is exposure to the different rates are appropriate for specific reasons. Helmets, seat-belts and child-restraints have an important role in death and injury prevention. Leading causes of death for young people aged 10-14 years were road injuries and people aged 15-19 years and 20-24-years it was transport injuries, Comparison groups were classified according to protective equipment use. Logistic regression and generalized liner models describe the probability of safety equipment use, head trauma, higher injury severity score and permanent sequel classified in road traffic injury.

II. EXISTING SYSTEM

However, as presented in the previous chapter, when studying road traffic fatal accident data. [6] if the analysis of roadway traffic data is critical to find out variables that are closely related to fatal accidents. We apply the statistical analysis and use data mining algorithm. Many attributes like collision manner, weather, surface condition, light condition, and drunk driver relationship between fatal rate were investigated. Certain safety driving suggestions were made based on statistics, association rules, classification model, and clusters obtained. A survey and detection of road safety aspects of national highway in India. [11] as we studied main sources of traffic crash data at the national level are the annual reports published by the National Crime Record Bureau (Ministry of Home Affairs) and the annual publication of the Ministry of Road Transport & Highways titled Road Accidents in India. The basic information both reports come from all the police station in the country reported to them. The Global Burden of Disease (GBD) study estimates that there were 264,000 (95% CI: 214,000-321,000) deaths in India in 2013 almost twice the deaths reported by traffic police. If the safety measures 80% of road accidents are attributed to the human error of momentary judgment (primarily drivers and vulnerable road users i.e. pedestrians, cyclist, rickshaw/cart pullers) suggested that the analysis of road transport and highways.

III. PROPOSED SYSTEM

The proposed work is analysis of real time data for road traffic accident (RTA). To overcome the existing problem of data missing and produce inefficient results also. To find further classification of data are categorized for different factors. The statistical analysis shows that the number accident in a selected city during the year and the causes of road traffic accident. Thus we have causes of accident records with age and gender information, the type of vehicle and the accident severity moreover in some city. This analysis to predict the accuracy of RTA given in the graphical representation.

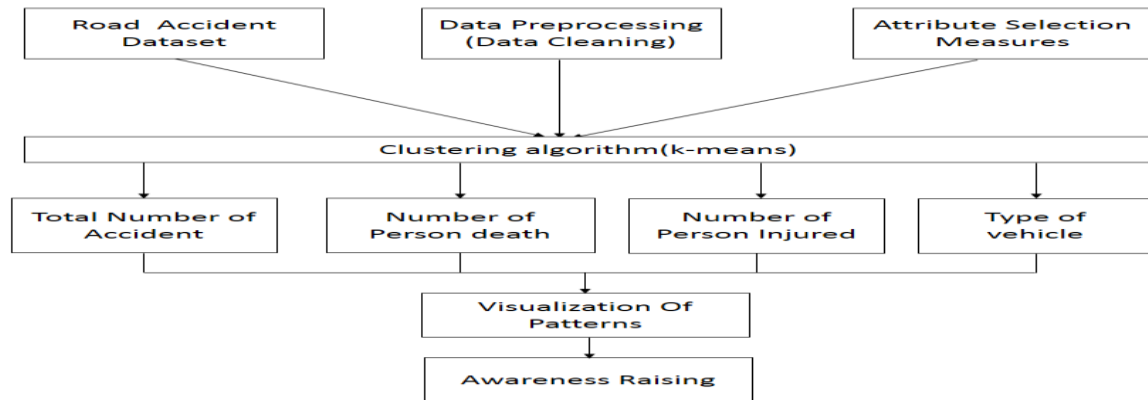


Fig1. Architecture Diagram of the proposed system

Accident dataset: A road traffic accident Data are collected is the systematic approach to gathering and measuring information from a variety of sources and repository to get a complete and accurate picture of an area of interest. The various attributes are included total accident, injuries, vehicle etc. **Data pre-processing:** In my work we have taken the accident RTA data set using a data mining technique that involves the raw data transformed my understandable format. We have to consider the accident dataset in a data inconsistent, incomplete, data missing, and is likely to contain many errors Data pre-processing is a proven method of resolving such issues. **Attribute selection measures:** A selection of an attribute in the road traffic accident dataset heuristic for selecting the splitting criterion that “best” separates a given year for data partition. **Naïve Bayes classification:** Classification in data mining methodology aims at constructing a model (classifier) from a training data set that can be used to classify records of unknown class labels. Naïve Bayes is the technique for used to based on the probability of model to incorporate strong independence assumption of each variable. **Visualization of patterns:** If a data are after the classification of Bayes theorem to predict the set of result it helps to people can understand the significance of data by placing it in a visual context. **Awareness rising:** we have Finally conclude the result of statistics for road traffic accident we may consider the environmental factor like, fatal accident, injuries, severity and so on. To make an analysis to generate the report for to give the awareness to the people.

IV. ANALYSIS AND RESULT

The training dataset collected from various repositories, then they have been cleaned and taken for data analytics. The output can be achieved using the IDE (R-studio) software. The report has been generated by the graphical representation of RStudio.

1. A measure of road traffic accident by age and gender for given year.

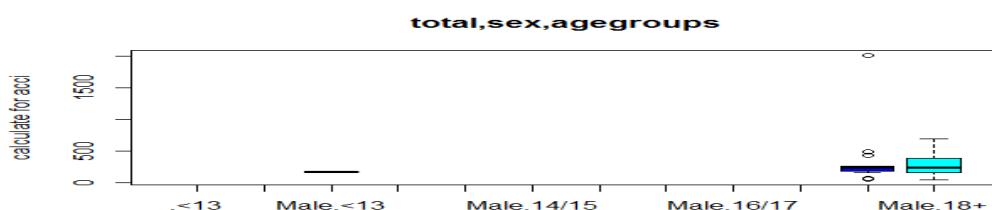


Figure 2: Accident ratio 2011

The figure 2, shows status of accident among different age groups of female and male members. In female group, between below 18 age group have minimum number of accidents (<18). Between above 18 age group have highest number of accidents (>18). In male group members, the below >13 age groups have minor accidents.

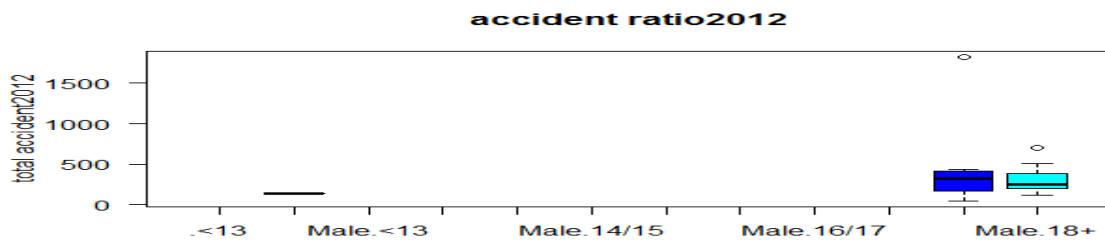


Figure 3: Accident ratio 2012

The figure 3, shows status of accident among different age groups of female and male members. In female group, between below 18 age group have highest number of accidents (>18). Between above 18 age group have minimum number of accidents (<18). In female group members, the below >13 age groups have minor accidents.

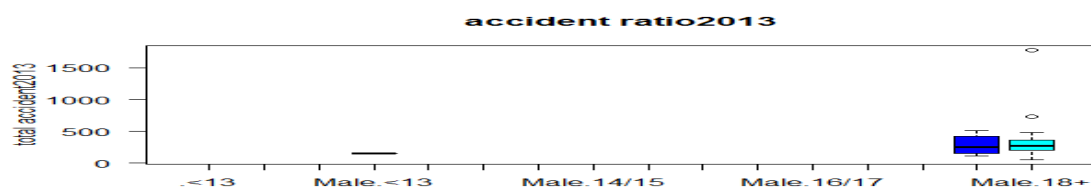


Figure 4: Accident ratio 2013

The figure 4, shows status of accident among different age groups of female and male members. In female group, between below 18 age group have highest number of accidents (>18). Between above 18 age group have minimum number of accidents (<18). In male group members, the below >13 age groups have minor accidents.

2. We have predicted and created a bar chart for the accident, happened for both males and females by all kinds of vehicles (Car, Truck, Bikes etc.)

The following bar- charts show the different kind of vehicle accidents among female and male separately in India on 2011, 2012.

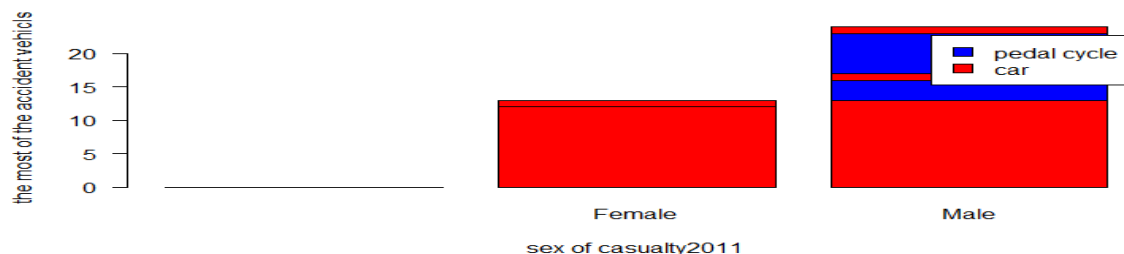


Figure 5: Causes of vehicle accidents (2011)

The 2011 Chart describes status of different type of vehicle accident among female and male members. In male group, have more accidents in a vehicle of car. In female group members, between major accidents in a vehicle of car, and in male group have minimum number of accident in a pedal cycle.

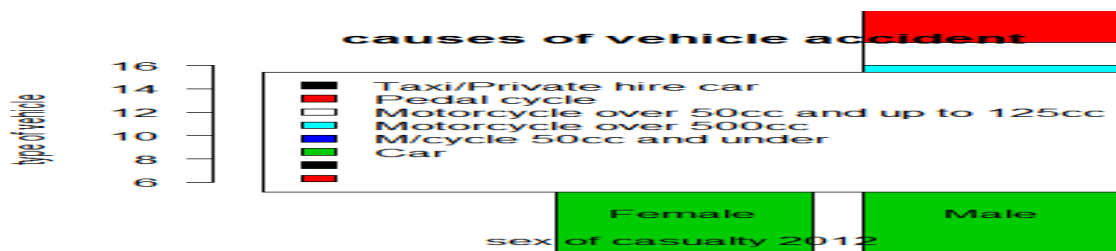


Figure 6: Causes of accidents (2012)

The 2012 Chart describes status of different type of vehicle accident among female and male members. In female group, have more accidents in a vehicle of car. In male group members, between major accidents in a vehicle of car, and in male group have minimum number of accident in a pedal cycle, motor cycle 500cc, motor cycle 50cc and up to 125cc.

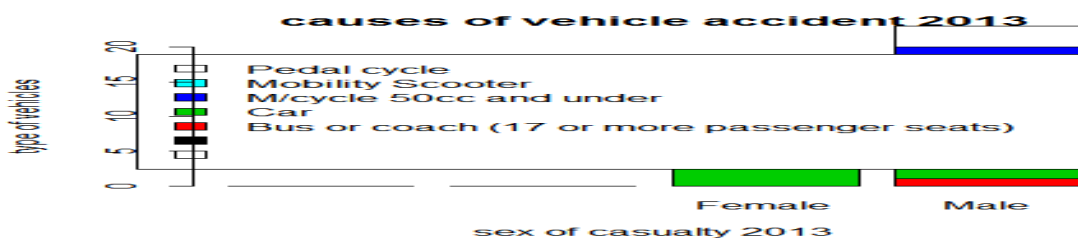


Figure 6: Causes of accidents (2013)

The 2013 Chart describes causes of different type of vehicle accident among female and male members. In male group, have more accidents in a vehicle of car. In female group members, between major accidents in a vehicle of car, and in male group have minimum number of accident in a pedal cycle, Bus or coach (17 more passenger seat), m/cycle 50cc and under.

2. I have predicted and created a boxplot on different kinds of age peoples.

The following boxplot shows the analysis of accidents among different kind of vehicles among different age groups in India on 2012 and 2013. If all y axis represents the total fatal accident in the year 2011-2013 respectively, if all x axis represents the all kind of vehicle accident happened among different age group

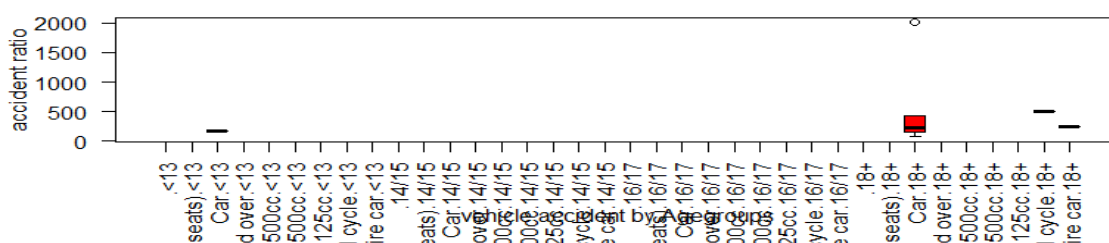


Figure 7: A vehicle accident in different age group (2011)

The figure 7, shows the visual report describes status of accident among different type of vehicles among different age groups members. In above 18 age group have more accidents (18+) in a vehicle of car. Between (>18) age group have major accidents in pedal cycle members, and between (>18) age groups have medium accident in vehicle of taxi/private hire car. between (<13) age group accidents have car.

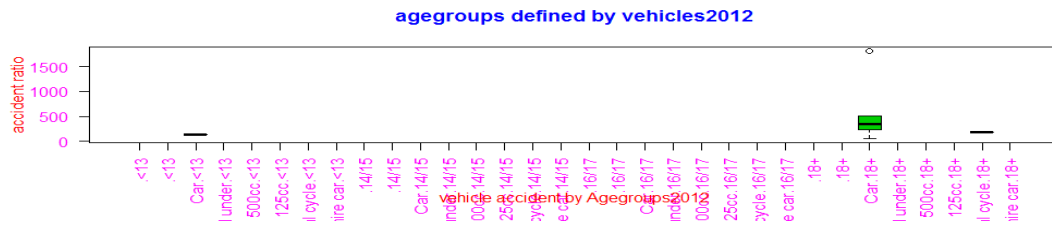


Figure 8: A vehicle accidents in different age group(2012)

The figure 8, shows the visual report describes status of accident among different type of vehicles among different age groups members. In above 18 age group have more accidents (18+) in a vehicle of car. Between (>18) age group have major accidents in pedal cycle members, and between (<13) age groups have medium accident in a car.

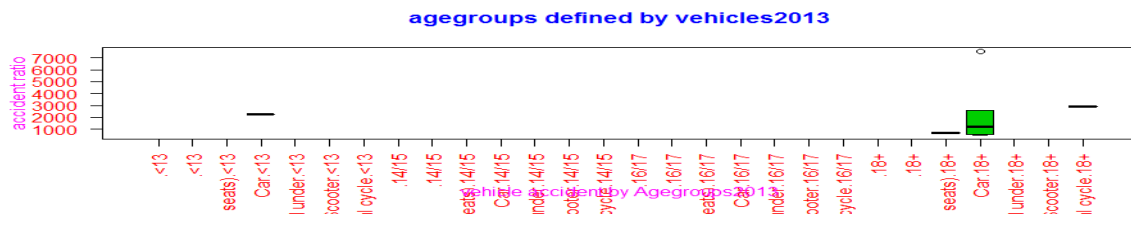
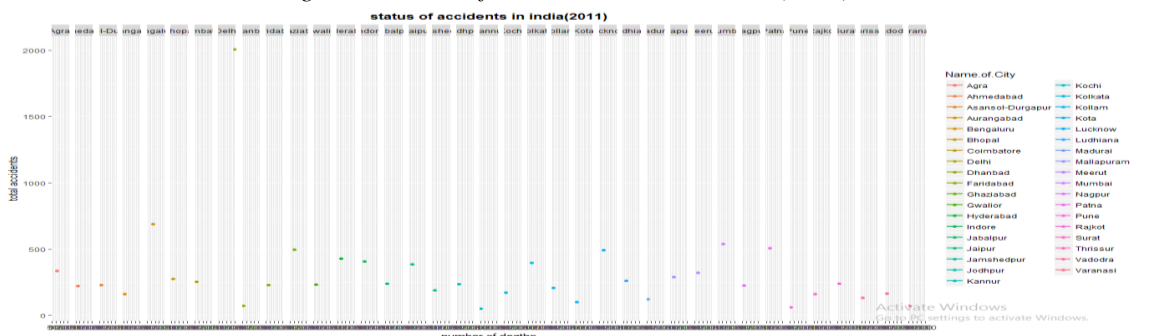


Figure 9: A vehicle accidents in different age group (2013)

The figure 9, shows the visual report describes status of accident among different type of vehicles among different age groups members. In above 18 age group have more accidents (18+) in a vehicle of car. Between (>18) age group have major accidents In pedal cycle members, between (<18) age groups have medium accident vehicle Bus or coach (17 or more passenger seats) . between (<13) age group have accident in car.

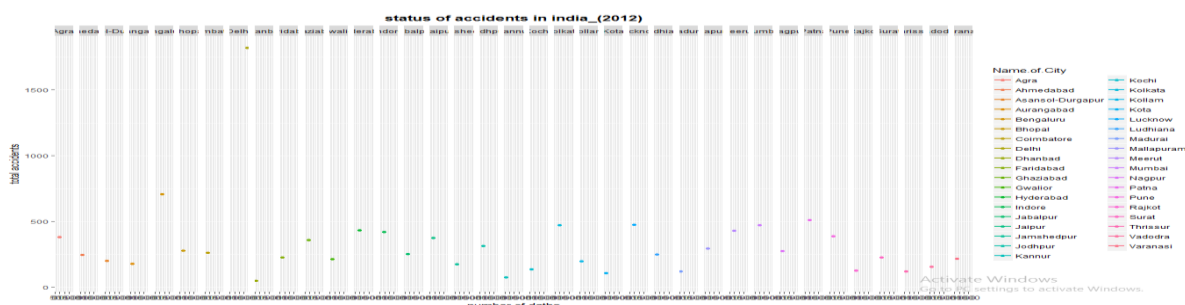
Figure 10: status of accidents in various cities (2011)



3. Accident death in various cities

The figure 10, shows the visual report of total number of accident deaths happened in various cities over all India in the year 2011. It describes the highest number of accident deaths happened in Delhi (>2000). Medium number of accident deaths happened in Bangalore(>600). Minimum number of accident deaths happened in Dhanbad, Kannur and Pune (<100).

Figure 11: status of accidents in various cities (2012)



The figure 11, shows the visual report of total number of accident deaths happened in various cities over all India in the year 2012. It describes the highest number of accident deaths happened in Delhi (>500). Medium number of accident deaths happened in Bangaluru(>600). Minimum number of accident deaths happened in Dhanbad, Kannur(<50).

4. Major, Minor and Non-injuries accident analysis

The following scatter- charts shows the analysis of major, minor and non-injuries accident happened among various cities in India on 2014 and 2015. In all Y axis represents major, minor and non-injuries respectively, and all X axis represents total numbers of accident happened among various cities.

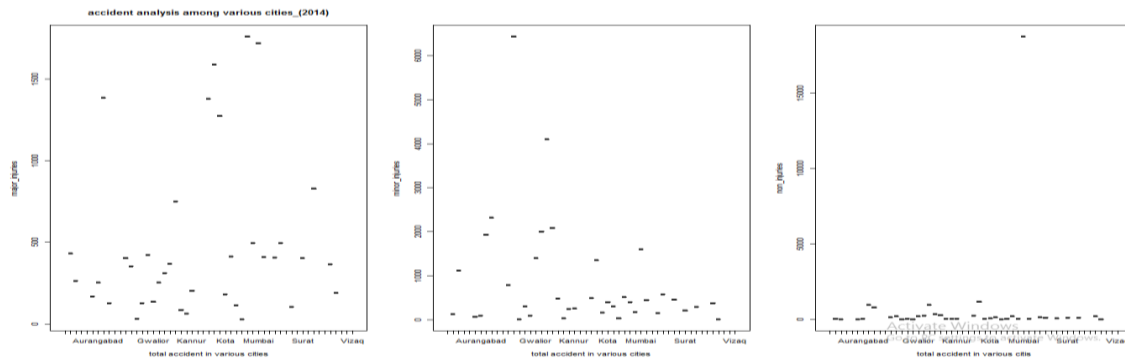


Figure 12: Accident analysis among various cities (2014)

The above Scatter- chart1 shows major injuries ,scatter-chart2 shows minor injuries and scetter-chart3 shows non-injuries happened among various cities in India on 2014 respectively. In Chart1 Kota and Mumbai cities have more number of injuries. In Chart2 Gwalior and Kannur cities have more number of Minor Injuries. In Chart3 Mumbai and Surat cities have more number of minor injuries

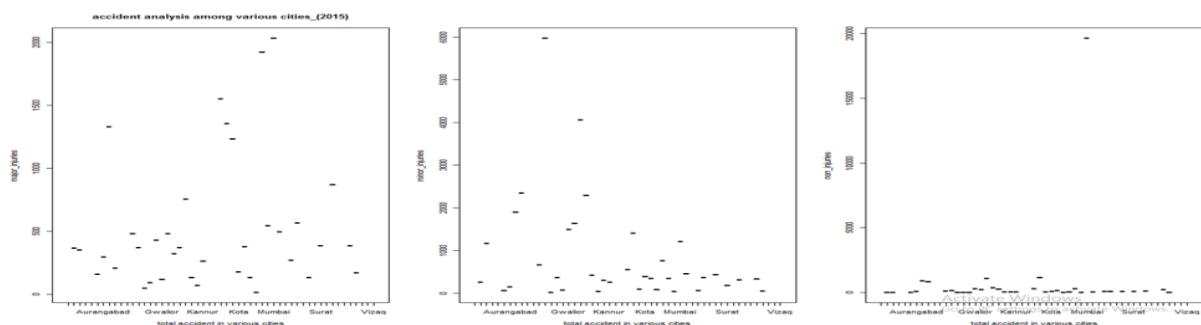


Figure 13: Accident analysis among various cities(2015)

The above Scatter- chart1 shows major injuries ,scatter-chart2 shows minor injuries and scetter-chart3 shows non-injuries happened among various cities in India on 2015 respectively. In Chart1 Kota and Mumbai have more number of major injuries. In Chart2 Gwalior and Kannur cities have more number of Minor Injuries. In Chart3 Mumbai and Surat cities have more number of non- injuries.

5. Accident injuries in age_group

The following scatter- charts show the Status of accident injuries among different age groups among female and male separately in India on 2012,2013,2014 and 2015 .

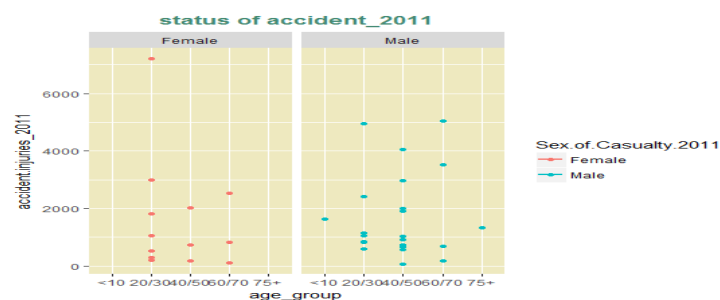


Figure 14: Status of accidents in (2011)

The 2012 Chart describes the status of accident among different age groups of female and male members. In female group, between 20-30 age group have more injuries (>6000). Between 60-70 age group have medium injuries. In male group members, the below <10 and between 60-70 age groups have major injuries (>4000), between 40-50 age group have minimum injuries.

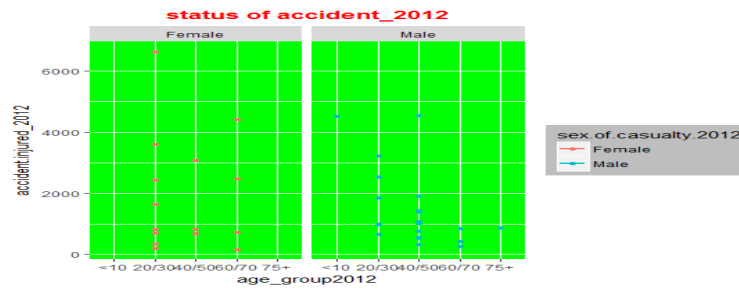


Figure 15: Status of accidents in (2012)

The 2012 Chart describes the status of accident among different age groups of female and male members. In female group, between 20-30 age group have more injuries (>6000). Between 60-70 age group have medium injuries. In male group members, the below <10 and between 40-5- age groups have major injuries (>4000), between 60-70 age group have minimum injuries.

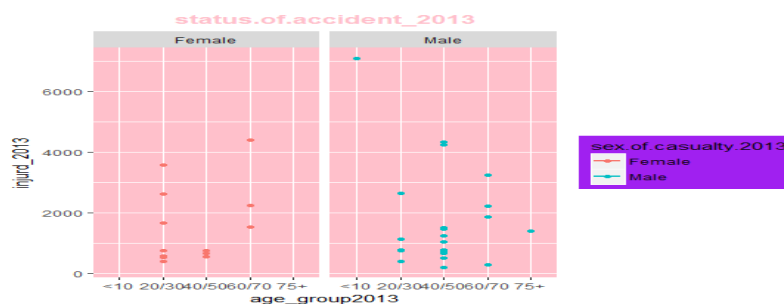


Figure 16: Status of accidents in (2013)

The 2013 Chart describes status of accident among different age groups of female and male members. In female group, between 60-70 age group have more injuries (>4000). Between 20-30 age group have medium injuries. In male group members, the below (<10) age groups have major injuries (>6000), between 40-50 age group have minimum injuries.

VI.CONCLUSION

Thus we have studied various methods and tools used to measure and predicted for road traffic accident in a dataset. This term has to be analysis for the selected city in the country of India and moreover the vehicle population in our country. It could be consider for the RTA dataset higher fatal rate in the year and stats that the city. Through the task performed, we realized that data seems never to be enough to make a strong decision. If more data, like person killed, injured, type of vehicle, gender and so on are available, more test could be performed thus more suggestion could be made from the data.

REFERENCE

- [1] [Www.safaribooksonline.com/library/view/getting-started-with/9781449314798/ch01.html](http://www.safaribooksonline.com/library/view/getting-started-with/9781449314798/ch01.html)
- [2] Ajit goswami* and ripunjoy sonowal division of epidemiology and nutrition, regional medical research centre, (indian council of medical research), n.e. region, dibrugarh, assam, india.
- [3] A statistical analysis of road traffic accidents in dibrugarh city assam India published in researchgate.net
- [4] Jorge gonzalez, solution engineer oct20,2017.
- [5] www.researchgate.net/profile/sridhar_seshadri2/project/safer-roads-safer-india
- [6] <https://www.bestcurrentaffairs.com/report-road-accidents-india-released/>

- [7] geethabaidepartment of computer science and engineering bangalore. .mariana antunes morgado,
department of pediatrics, santa maria hospital, portugal, Tel: 00351918275875; e-mail:
morgadoamariana@gmail.com.
- [8] government of india ministry of road transport & highways Transport research wing.
- [10] liling li, sharad shrestha, gongzhu hu department of computer science central Michigan university, usa (li8l,
shres1s, hu1g)@cmich.edu.
- [11] pankaj dahiya m.tech. Student, dept. Of civil engg., cbs group of institutions Jhajjar,haryana.
- [12] bhattacharya, s., alberini, a., & cropper, m. L. (2006). *The value of mortality risk Reductions in delhi, india*
(publication., from university of maryland and world bank:
- [13] george yannis, professorwww.nrso.ntua.gr/geyannis.
- [14] Péter holló,habil, d. Sc. Head of dept. For road safety university professor and traffic Engineering
széchenyi istváninstitute for transport sciences university,
- [15] Liling li, sharad shrestha, gongzhu hu department of computer science central Michigan university, usa
(li8l, shres1s, hu1g)@cmich.edu.
- [16] Ramya v#1department of information science and engineering,dayananda sagar College of engineering
bengaluru karnataka 560078. *1ramy.viju@gmail.com*.