

# Expressway Traffic Accidents Involving Human Injuries and Fatalities in Ethiopia: Negative Binomial Regression Model

Yesuf Abdela Mustefa (✉ [yesufabdela@yahoo.com](mailto:yesufabdela@yahoo.com))

Technology and Innovation Institute of Ethiopia <https://orcid.org/0000-0003-2184-0845>

Addis Belayhun

Technology and Innovation Institute of Ethiopia

---

## Short Report

**Keywords:** Expressway, Traffic, Accidents, Injuries, Fatalities, Negative Binomial

**Posted Date:** April 15th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-151181/v2>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Expressway Traffic Accidents Involving Human Injuries and Fatalities in Ethiopia: Negative Binomial Regression Model

Yesuf Abdela Mustefa<sup>1\*</sup> and Addis Belayhun<sup>2</sup>

<sup>1</sup>Computing and Data Analytics Directorate, Technology and Innovation Institute, Addis Ababa, Ethiopia

<sup>2</sup> Computing and Data Analytics Directorate, Technology and Innovation Institute, Addis Ababa, Ethiopia

\*Corresponding author name and email address: Yesuf Abdela Mustefa - yesufabdel777@gmail.com

## Abstract

**Background:** Road traffic accident is a major public health as well as economic challenge that rated the eighth leading cause of death. The severity became higher in developing countries. Ethiopian is among the most confronted countries in the world. We utilized the Ethiopian Toll Roads Enterprise data to provide insights and model significant determinants of accidents involving injuries and fatalities. Besides utilizing recent dataset, we applied the most appropriate but forwent statistical model. Moreover, we examined the significance of the effects of drivers' age and gender that have not been the cases in the literatures.

**Methods:** We made descriptive insights available on the basis of graphs from integrated traffic accident and flow datasets. We tested for the presence of over-dispersion in a total of 1824 observations of accident data recorded from September, 2014 to December, 2019 for inferential analysis. Finally, we modeled the effects of significant variables on the number of injuries using the negative binomial regression model.

**Results:** we found that the number of injuries in accidents were significantly determined by type of vehicles, ownership status of vehicles, accident time weather condition, driver-vehicle relationship, drivers' level of education, and drivers' age.

**Conclusions:** Heavy trucks were more likely to cause more number of injuries than medium or small vehicles. Hot and windy weather conditions were associated with higher probability of the number of injuries. The likelihood of the number of injuries were lower when drivers are owner of the vehicle; drivers level of education is above secondary school; and the age of the driver is between 18 and 23 years old. Moreover, due concern needs to be given for traffic road rules.

**Keywords:** Expressway; Traffic; Accidents; Injuries; Fatalities; Negative Binomial.

## Background

Road traffic accident (RTA) is a major global concern (Abegaz T, 2019). It is among the leading causes of death. Considerable number of people die around the world every year and several millions of people leave with non-fatal injuries as a result of RTA (WHO, 2018). Furthermore, RTA seriously impacts economies since it results in loss of productivity of those killed as well as disabled and incurs treatment costs for injured. According to (WHO, 2018), it costs significant amount of annual gross domestic product of countries. Thus, RTA affects public

health and development in more general terms (Abegaz T, 2019).

Recent findings revealed that the impact and rate of RTA are relatively more and increasing over time in low and middle income countries (Abegaz T, 2019; WHO, 2018). Especially, the highest annual rate of RTA fatalities are recorded in Africa and it is even expected to rise with the growing economy of the region (Abegaz T, 2019; WHO, 2018).

Despite the remarkable enhancements that has been made in public health status, RTA is a major

public health problem in Ethiopia as one of the countries in Africa (Anteneh & Endris, 2020; Jembere GB, 2018). According to (WHO, 2018), the rate of RTA fatality was estimated to be 27 per 100, 000 population which indicates the gravity of the problem. The road safety record of the Ethiopian expressway is also alarming the severity of the situation and calling for an integral effort of all pertinent stakeholders to reverse the trend (Deme & Bari, 2018). The expressway often called Addis Ababa-Adama Expressway is the first modern road in Ethiopia and East Africa as well which officially opened for traffic in September 2014 (Abera, 2019). The road has technologically advanced safety management system aiming to reduce traffic accident. However, traffic accidents occur frequently (Abera, 2019). Records of the Ethiopian Toll Roads Enterprise indicated that, about 672 traffic accidents were reported in one and half year since the expressway was open for traffic (Deme & Bari, 2018). The accident raised to 1137 in three years and eight months since the launch of the road (Abera, 2019). Moreover, increase in the number of vehicles is contributing to increased number of accidents (Kuleno, et.al., 2018). Data from September, 2014 to December, 2019 that we obtained from the enterprise also shows that the absolute magnitudes of traffic flows and accidents are increasing over time. However, to the best of the researchers' knowledge, no studies have found on the basis of recent datasets of the enterprise. Most importantly, the effects of drivers' age and sex on accidents involving injuries and fatalities have not been investigated. Thus, we utilized forwent but appropriate statistical method of the arena specifically, the negative binomial regression (NBR) model in order to determine significant factors contributing to traffic accidents in the expressway. Besides, we explored the rates of traffic flows and accidents over time in association with various potential attributes so

that appropriate measures can be taken by stake holders based on the insights.

## Method

The data was obtained from Ethiopian Toll Roads Enterprise. We analyzed a total of 1824 observations of accident data recorded from September, 2014 to December, 2019. The enterprise's real time system recorded 35387231 traffic flow data with in the specified period. We integrated the two datasets based on key variables to explore the yearly variations and trends of accident rates.

The accident dataset contained 13 potential attributes which were supposed to determine total number injuries in an accident namely, vehicle type; number of service years of vehicle; type of injury ;accident time; cause of accident; road geometry; accident time weather condition; light condition; driver-vehicle relation; driving experience; drivers' educational level; age and gender of driver. We used descriptive charts and graphs to extract insights from available attributes of the data. Finally, we used the NBR to model significant determinants of the number of injuries. We used the NBR model because it captures the over-dispersion in the data better than the Poisson regression model (Zeileis, et. al., 2008). Moreover, no excess zeros were found in the data. As a result, the negative binomial regression model was best fit to the data among all other count regression models.

### The NBR Model

The NBR model for count data belongs to the family of generalized linear models in which the link function is logarithm of the mean number of injuries total injuries in accident  $\log(\mu)$  (Zeileis, et. al., 2008). Total number of injuries in an accident was defined to be the number of light, serious, and fatal injuries in an accident. The NBR model based on significant variables in this study is given as (equation 1). Variables that were

not significant in determining the total number of injuries were excluded from the final NBR model.

$$\begin{aligned} \log(\mu) = & \text{Intercept} + \beta_{1i} \text{Vehicle Type}_i \\ & + \beta_{2i} \text{Vehicle Ownership}_i + \\ & \beta_{3i} \text{Weather Condition}_i + \\ & \beta_{4i} \text{Driver Vehicle Relationship}_i \\ & + \beta_{5i} \text{Driver's Education Level}_i \\ & + \beta_{6i} \text{Driver's Age}_i \end{aligned} \quad (1)$$

Where: the subscript  $i$  indicates categories of the corresponding variable; the  $\beta$  s' are coefficients which are additive effects on the log of average number of total injuries of categories of the corresponding variable as compared to the reference category;  $\mu$  is the average total number of injuries in accident.

We also checked the appropriateness of the NBR model by testing for the presence of over-dispersion in the data.

## Results and Discussions

### Insights from Descriptive Analysis

The percentage of distinct vehicles passing through the expressway as well as their flows were increasing over time. However, the changes of percentages between the distinct vehicles and their flows differed. The average yearly changes in percentages of distinct vehicles was 8% which is approximately double as compared to the average yearly changes in vehicles flow which was found to be 4.2% [Figure 1].

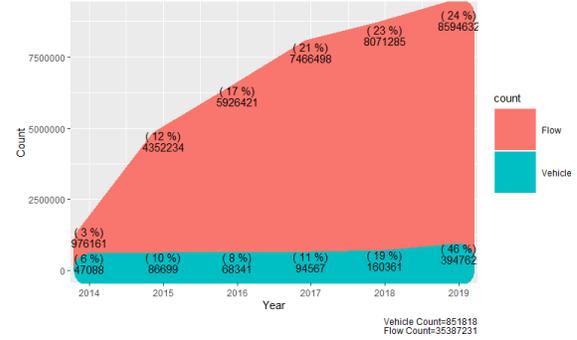


Figure 1: Relative yearly variations in distinct vehicles and their flows

On the other hand, the average yearly changes of percentages in accident is approximately 2% which is one half as compared to vehicles flow percentage. However, the average percentage of accident is overestimated by rapid increases in the first three years. The level of accident began to decrease from 2018 forward [Figure 2]. Thus, the stake holding body of the expressway should strengthen its practice in 2019 besides taking further actions required for lowering the level of accident.

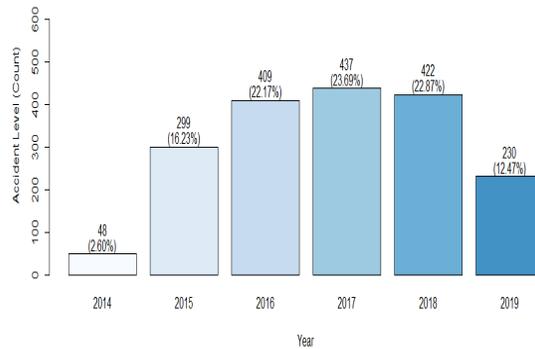


Figure 2: Yearly percentages of accident

In order to eliminate the effect of increasing flow on accident, we calculated the rates of accidents per 100K flows for every single year. Although the average absolute changes in percentages of accident was positive indicating an increase, the rate of accidents in reference to vehicles flow was decreasing [Figure 3].

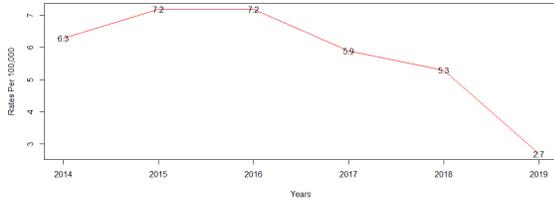


Figure 3: The rates of accidents per 100K

We explored the levels of injury caused from traffic accidents. Of the total 1824 traffic accidents, 1181(65%) of accidents caused various injuries. Among the casualties, (130)11% were dead and the remaining (324) 27% and (727) 62% suffered respectively from serious and light injuries [Figure 4]. According to (Abegaz T, 2019), the rate of RTA-related injury was 163 (95% confidence interval (CI): 136–195) per 100,000 population.

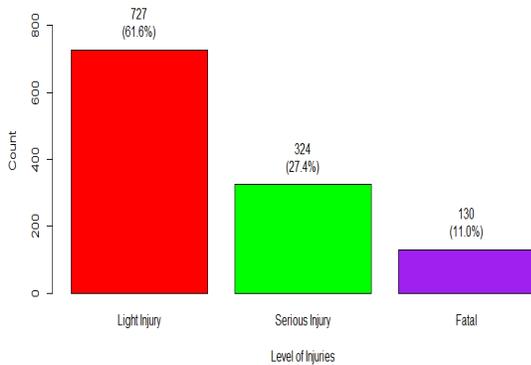


Figure 4: Level of injuries caused from traffic accidents

Not only the rates of accidents but also, the rates of injury levels were decreasing [Figure 5].

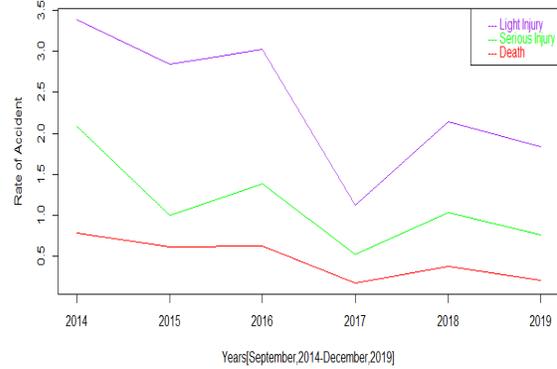


Figure 5: The rates of injury levels

Accidents varied over hours within days. The rate of accident was higher at night than day times. There was 7 accidents in every 100K flows at the early night and 9 at the late night. The lower rate observed at day times before noon and it was 4 in every 100K flows. The after noon takes the third place in which 5 accidents occurred in 100K flows [Figure 6]. However, according to (Honeln & Wuletaw, 2020; Deme & Bari, 2018; Abera, 2019), the absolute magnitude of accidents is higher in the afternoon. This study also confirmed the same regarding the absolute number of accidents but different relative to the number of traffic flows. The number of accidents depends on the number of traffic flows (Abera, 2019). Therefore, regardless of the manner of interpretations, indications of results of the studies coincide.

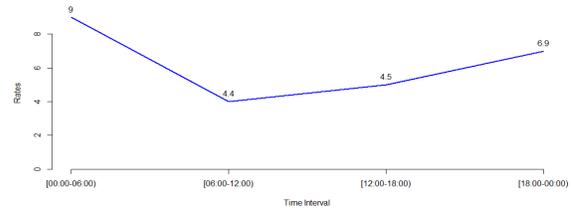


Figure 6: Time interval rates of accident per 100K flows

Moreover, drivers of heavy vehicles should be more conscious in their journey since heavy sized vehicles were more attributing to accidents than small sized vehicles. In 100K flows of heavy vehicles, there were 8 accidents. On the other

hand, mini sized buses were less exposed than automobiles [Figure 7]. However, in absolute count automobiles were most frequent in both flow and accidents (Deme & Bari, 2018).

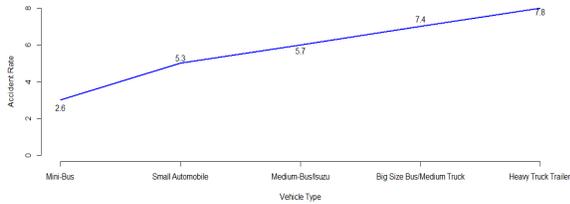


Figure 7: Vehicle type variation of accidents per 100K flows

Besides vehicle type specific concern, unethical driving or disobeying road traffic rules is another issue which accounted for 41% of the total accidents. Speed above or below the specified limit is one of unethical driving. Not only speed but also speed change affects the occurrence of accidents (Tanishita & Wee, 2017). A study also confirmed that major causes of traffic accidents were associated with unethical driving (problem with over speed, sight distances, and phone usage during driving, drug usage and etc.) behaviors that caused 59.8% of the total accidents (Deme & Bari, 2018). Another 30% caused from technical problems on vehicles; 4% from fatigue; approximately 1% because of animals; and only small portion 0.2% of the accident was caused from flood. However, flood holding areas of the road should be maintained in order to lower the rate of accidents. The reasons for the remaining 24% of accidents were unspecified [Figure 8].

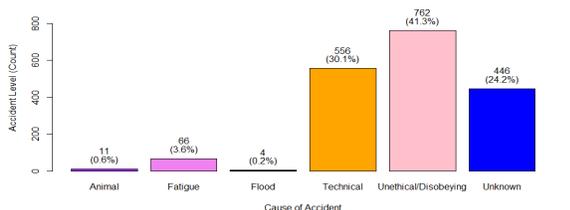


Figure 8: Cause of accidents

Controlling mechanism should be established to enforce traffic road rules. And also, drivers needs

to develop the habit of checking vehicles before journey. So that they can lower the possibility of accidents caused from technical problems of vehicles. Stake holders should build fences where animals are unexpectedly entering the road. Despite the small percentage, the rate of driving without resting enough was increasing [Figure 9]. Thus, drivers' status checking mechanism should be established at initial plazas.

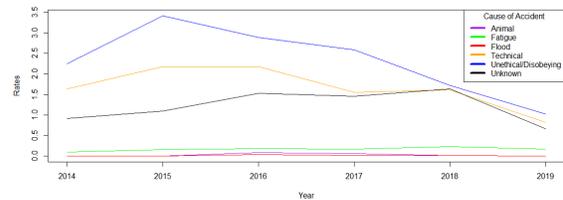


Figure 9: Cause specific accident rates

All in all, if animal intrusion, fatigue driving, flood, vehicles' technical problem, and unethical or disobeying traffic rules were the only factors contributing to the occurrence of accident, avoiding animal intrusion from the expressway possibly decreases accidents approximately by 1% at the minimum [Figure 10]. It will approximately decrease approximately by 25% if all of the 24% unknown causes in [Figure 9] were animals. In the same manner, ethical driving, technical assurance, road maintenance, and avoiding restless driving result in (41%-66%), (30%-54%), (0%-24%), and (4%-28%) minimum and maximum decreases in accidents respectively [Figure 10].

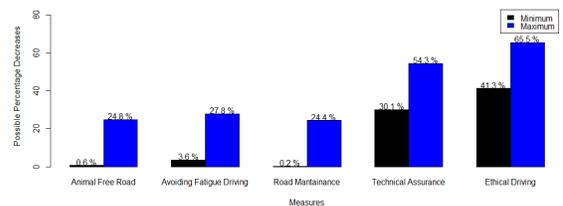


Figure 10: Possible measures to decrease accidents

We cross categorized accident causes by gender of drivers. Approximately, only 1% were female

drivers and 83% were males. The remaining 16% remained unspecified [Figure 11].

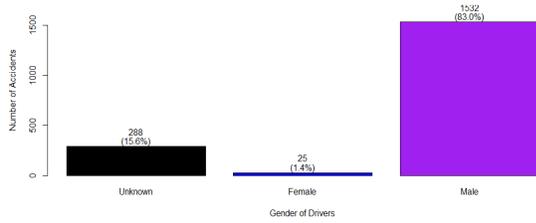


Figure 11: Gender of drivers

Although the females were small in number, of the total 25 females, we found that they were driving more unethical or disobeying relative to males. Thus, we recommend license providers should pay due attention [Figure 12]. However, gender was not significant in the final model.

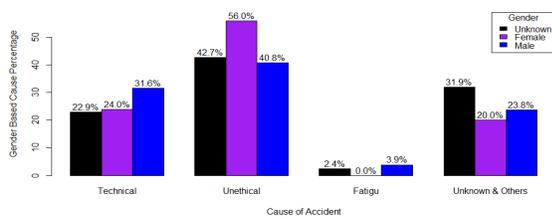


Figure 12: Cross classification of accident causes by gender

Other drivers' characteristics that we found significant was age of drivers. Most of the accidents were associated with drivers of young age [Figure 13]. This indicates that traffic accident affected productive man power of the country. Thus, the impact of traffic accident extended to economic challenges beyond public health issues (Abegaz T, 2019).

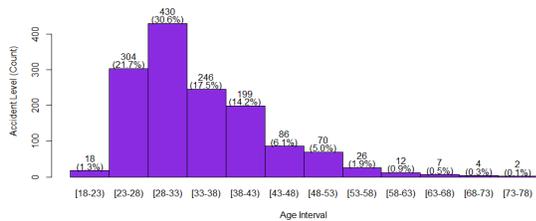


Figure 13: Drivers' age distribution

The percentage of accidents also varied with road geometry. Most (34%) of the accidents occurred on sloppy road [Figure 14].

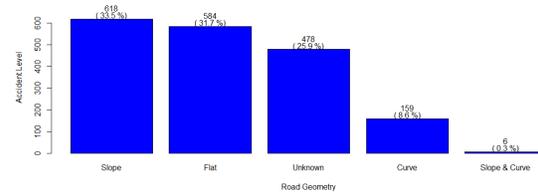


Figure 14: Accident road geometry

If the only factor contributing to the occurrence of accidents was geometry of the road, taking the most care and/or measures at slopes and curves would reduce 34%-60% and 9%-35% of the accidents [Figure 15].

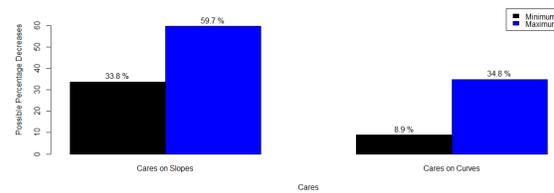


Figure 15: Percentages of accident decreases by taking cares at slopes and curves

We summarized accident time weather conditions. However, the longest duration throughout the year is normal in Ethiopia. Thus, our assessment on accident time weather conditions revealed that larger percentage of accidents occurred at normal weather condition [Figure 16].

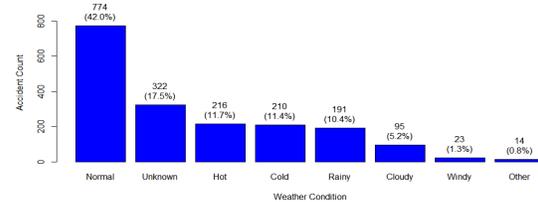


Figure 16: Accident time weather conditions

### Estimates of NBR Model Parameters

Before fitting the NBR model, we tested the presence of over dispersion in the data. The

estimate of the dispersion parameter was 3.33238 which clearly greater than the null value 1. The corresponding Z-value was 3.7485 that resulted in a P-value of 0.00008896 which is far lower than the level of significance (0.05) [Table 1]. Thus, we rejected the null hypothesis that states the dispersion is 1 and concluded that there is over-dispersion in the data at 5% level of significance and the NBR model more appropriate fit to the data than Poisson regression model.

Table 1: Over-dispersion test

Estimate of Dispersion	Z	P-Value
3.33238	3.7485	8.896e-05

We tested all attributes in the dataset that potentially expected to contribute to the occurrence of injuries. Accordingly, we found that vehicle type, ownership of vehicles, accident time weather condition, drivers' relation with vehicles, education level of drivers, and age of drivers were significant determinants of number of injuries [Table 2].

Table 2: Parameter estimates of NBR model

Factors	Categories	Estimate	SE	Pr(> z )
	Intercept	-2.511	1.061104	0.0179
<b>Vehicle Type (Ref: Unspecified)</b>	Big Size Bus	1.293	1.052125	0.2189
	Medium Truck (3Axel)			
	Heavy Truck Trailer(>=4 Axel)	1.386	1.046614	0.1854
	Medium Bus & Small Truck (2 Axel)	2.346	1.037234	0.0237
	Minibus	3.613	1.041141	0.0005
<b>Ownership (Ref: Company)</b>	Small Automobiles & 4WD	2.074	1.034441	0.0449
	Government	0.009	0.350768	0.9806
<b>Weather Condition (Ref: Normal)</b>	Private	0.499	0.230324	0.0303
	Unspecified	0.242	0.250862	0.3348
<b>Weather Condition (Ref: Normal)</b>	Hot	0.457	0.181940	0.0120
	Normal	0.078	0.137193	0.5676

<b>(Ref: Cloudy/Misty)</b>	Unspecified	-0.365	0.174080	0.0359
	Windy	0.172	0.505791	0.7340
<b>Driver-Vehicle Relation (Ref: Unspecified)</b>	Employee	-0.500	0.180237	0.0056
	Other	-0.285	0.234047	0.2237
<b>Drivers' Education Level of Education (Ref: Unspecified)</b>	Owner	-0.607	0.209179	0.0037
	Unspecified			
<b>Drivers' Education Level of Education (Ref: Unspecified)</b>	Above Secondary	-0.855	0.203963	2.78e-05
	Primary	-0.790	0.252983	0.0018
<b>Driver Age (Ref: Unspecified)</b>	Secondary	-0.490	0.176849	0.0056
	[18-23)	-1.193	0.694820	0.0860
	[23-28)	-0.496	0.200207	0.0133
	[28-33)	-0.149	0.179081	0.4070
	[33-38)	-0.475	0.208279	0.0227
	[38-43)	-0.417	0.219141	0.0571
	>=43	-0.169	0.219206	0.4401

Ref: Reference; Pr: Probability; SE: Standard Error

The number of injuries occurred in association with small automobiles & 4WD; min-buses; and medium sized buses & small trucks (2 axel) were significantly different from that of unspecified vehicle types. On the other hand, injuries as a result of heavy truck trailers (>= 4 axel); and big sized buses & medium trucks (3 axel) were not significantly different from the reference category. Thus, accident on small automobiles & 4WD; min-buses; and medium sized buses & small trucks (2 axel) resulted in significantly different number of injuries from accident on heavy truck trailers (>=4 axel); and big sized buses & medium trucks (3 axel) and it increases the log of average number of injuries by 2.074, 3.613, and 2.346 respectively as compared to accident in unspecified vehicle types holding the effects of other factors constant at 5% level of significance [Table 2]. The significance of vehicle type was confirmed in other study (Deme & Bari, 2018).

In the same manner of interpretation, the log of average number of injuries due to accidents on vehicles with private ownership significantly increases by 0.499 as compared to that of vehicles with company ownership. Accidents on vehicles of government and unspecified ownership were

not significantly different from vehicles of company ownership. Hot weather condition results in 0.457 increase. The log average number of injuries decrease more (by 0.607) when the driver is owner of the vehicle. Drivers' level of education significantly contributes to the number of injuries in accidents. Drivers' age intervals [23-28) and [33-38) were more likely associated with reduced number of injuries [Table 2].

## **Conclusions and Recommendations**

Although the rate of accidents were decreasing, stake holding body of the expressway should further work on decreasing the absolute number of accidents since the fatality due to traffic accident is considered to be man made disaster in Ethiopia (Mohammed, et.al., 2019). Highest frequency of accidents occurred as a result of unethical or disobeying driving. Thus, awareness need to be created in order to lower accidents. Besides, controlling mechanisms and traffic road rules enforcement are possible lowering measures. Moreover, immediate maintenance of broken roads, and preventing animals from entering the expressway positively contribute to accident controlling. The percentage of females driving unethical is higher than that of males. Therefore, more needs to be worked on females' awareness towards road traffic rules and license providing institutions should take further concern. The absolute flow and accident on small automobiles is maximum. However, the rate of accidents based on 100K flows indicated that heavy vehicles were more likely associated to accidents. Thus, truckers should be more conscious since heavy vehicles are more difficult to control. Further, due concern needs to be given for night time traffic movements. Nevertheless, it is desirable to take all possible actions in order to eliminate traffic accident as it is continued being public health and development issue.

In general vehicle type, ownership status of vehicle, accident time weather condition, driver-

vehicle relation, drivers' level of education, and drivers' age were significant determinants of the number of accidents involving injuries and fatalities at 5% level of significance. The rate of accident varied across the levels of the remaining variables. However, it was not significant according to the final NBR model.

## **List of Abbreviations**

RTA: Road Traffic Accident; NBR: Negative Binomial Regression

## **Declarations**

### **Ethics Approval and Consent to Participate**

Not Applicable.

### **Consent for Publication**

Not Applicable.

### **Availability of Data and Material**

The data analyzed during this study is stored at the database of Ethiopian Toll Roads Enterprise and it will be available on a reasonable request.

### **Competing Interests**

The authors declare that they have no competing interests.

### **Funding**

Not Applicable.

### **Authors' Contributions**

Both authors YAM and AB generated the idea, YAM as a corresponding author analyzed and interpreted the data, AB contributed his part on descriptive analysis of the data under the supervision of YAM. Both authors read and approved the final manuscript.

### **Acknowledgements**

We acknowledge the Ethiopian Toll Roads Enterprise for giving us the dataset and the big data analytics team in Technology and Innovation Institute of Ethiopia for helping us in cleaning the data.

## References

- Abegaz T, G. S. (2019). Magnitude of road traffic accident related injuries and fatalities in Ethiopia. *PLoS ONE*, *14*(1), 1-10. Retrieved from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0202240>
- Abera, M. A. (2019). Evaluation of road traffic accident in Addis Ababa –Adama Expressway. *International Journal of Advance Research, Ideas and Innovations in Technology*, *5*(5), 450-456. Retrieved from <http://www.IJARIT.com>
- Anteneh, A., & Endris, B. S. (2020). Injury related adult deaths in Addis Ababa, Ethiopia: analysis of data from verbal autopsy. *BMC Public Health*, *20*(926). Retrieved from <https://bmcpublihealth.biomedcentral.com/articles/10.1186/s12889-020-08944-7>
- Deme, D., & Bari, M. (2018). Traffic Accident Causes and Its Countermeasures on Addis Ababa-Adama Expressway. *Journal of Equity Science and Sustainable Development*, *2*(2), 13-23. Retrieved from <http://www.mwu.edu.et/journal-sites>
- Honelgn, A., & Wuletaw, T. (2020). Road traffic accident and associated factors among traumatized patients at the emergency department of University of Gondar Comprehensive Teaching and Referral Hospital. *PAMJ Clinical Medicine*, *4*(9). doi:10.11604/pamj-cm.2020.4.9.24507
- Jembere GB, Y. C. (2018). Decomposition of Ethiopian life expectancy by age and cause of mortality; 1990–2015. *PLoS ONE*, *13*(10).
- Kuleno, M. M., Denno, H. D., Abera, G., & Reddy, D. R. (2018). Quantifying Traffic Congestion by Studying Traffic Flow Characteristics in Wolaita Sodo Town, Ethiopia. *International Journal of Scientific & Engineering Research*, *9*(6), 109-120.
- Mohammed, A. A., Ambak, K., Mosa, A. M., & Syamsunur, D. (2019). A Review of Traffic Accidents and Related Practices Worldwide. *The Open Transportation Journal*, *13*, 65-83. doi:10.2174/1874447801913010065
- Tanishita, M., & Wee, B. V. (2017). Impact of Vehicle Speeds and Changes in Mean Speeds on Per Vehicle-Kilometer Traffic Accident Rates in Japan. *International Association of Traffic and Safety Sciences (IATSS)*, *41*(3), 107-112. doi:http://dx.doi.org/10.1016/j.iatssr.2016.09.003
- WHO. (2018). *Global Status Report on Road Safety*. Geneva: World Health Organization. Retrieved December 22, 2020, from World Health Organization: <https://www.who.int/publications/i/item/9789241565684>
- Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression Models for Count Data in R. *Journal of Statistical Software*, 1-25. doi:10.18637/jss.v027.i08 · Source: DOAJ

# Figures

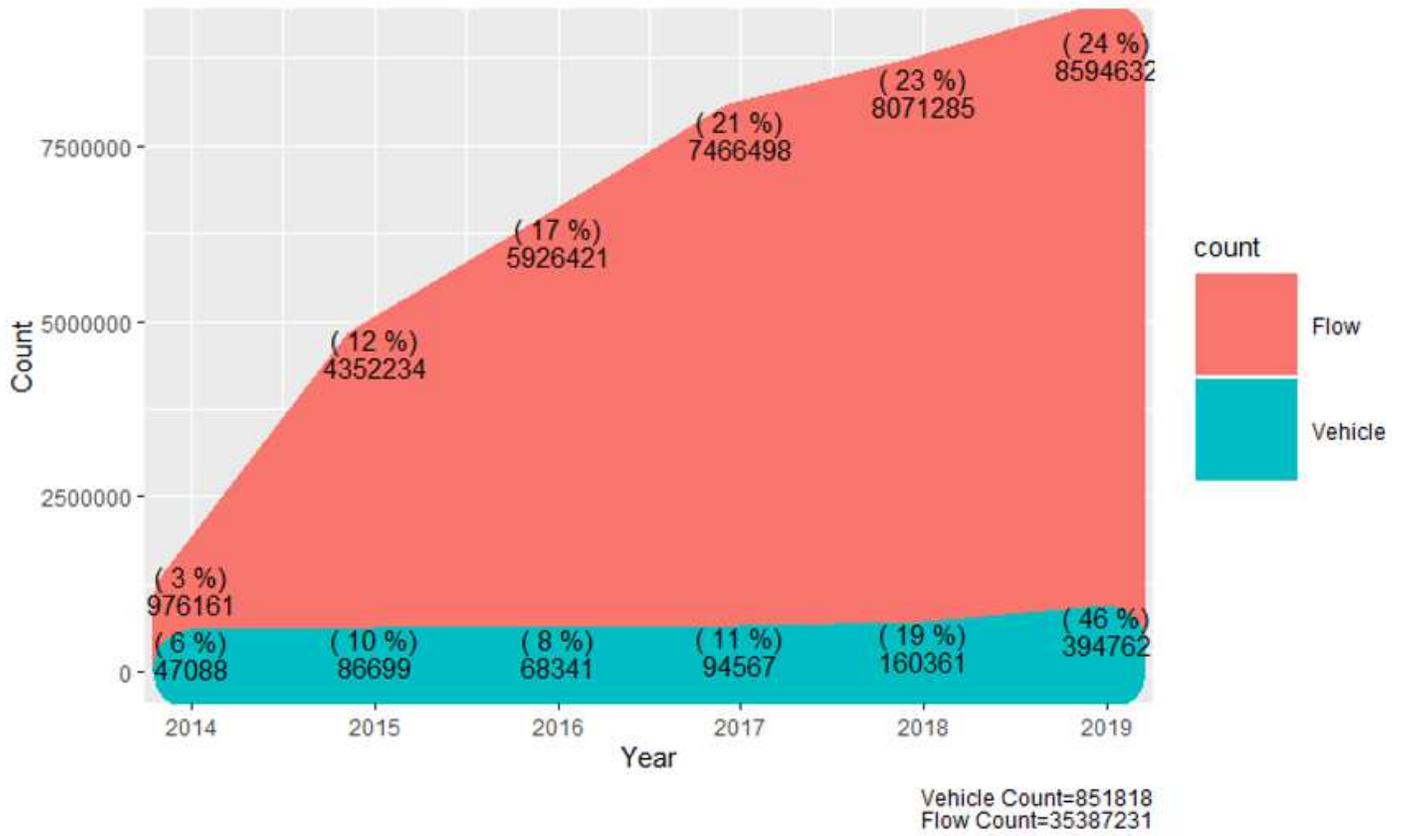
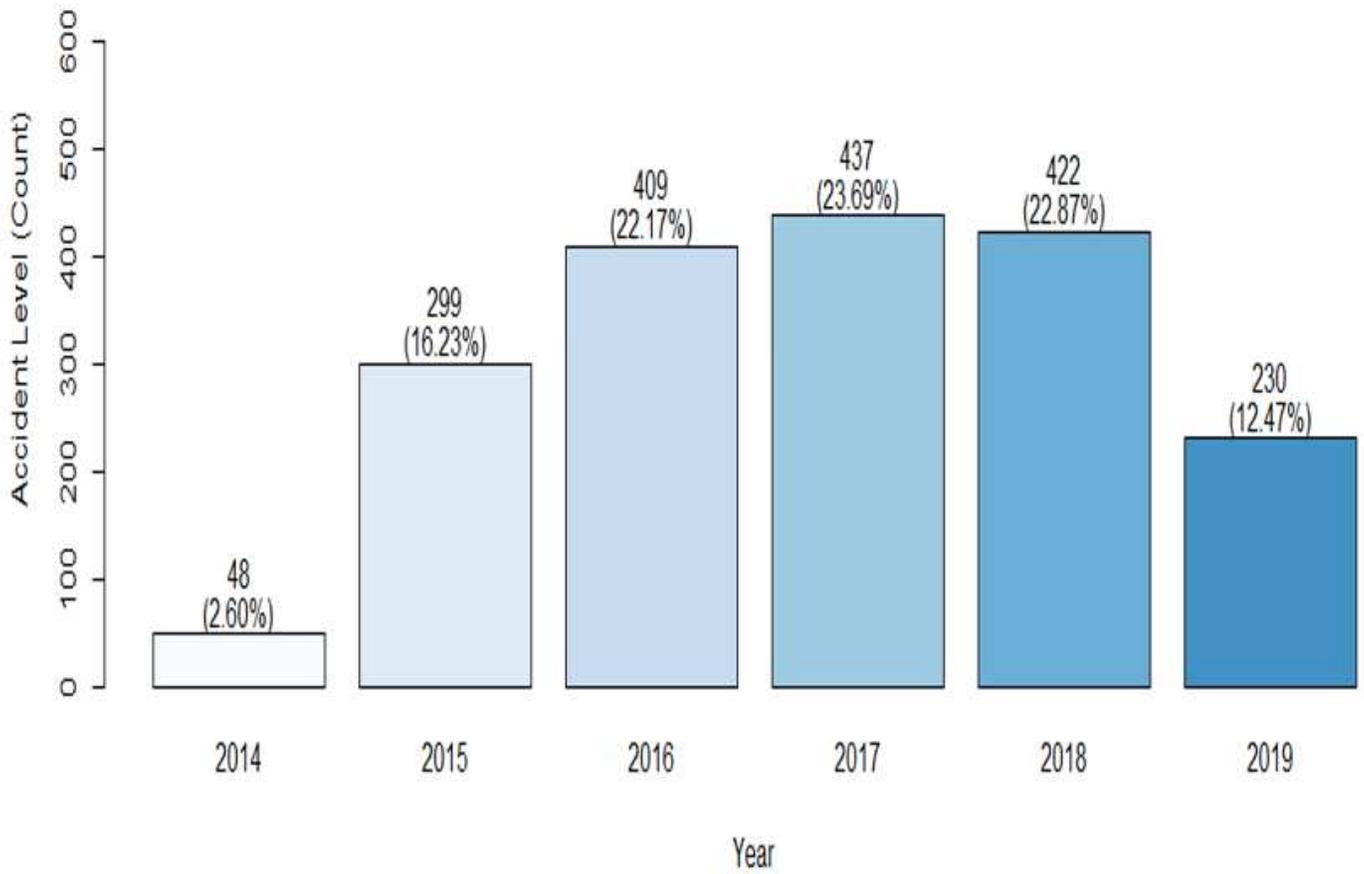


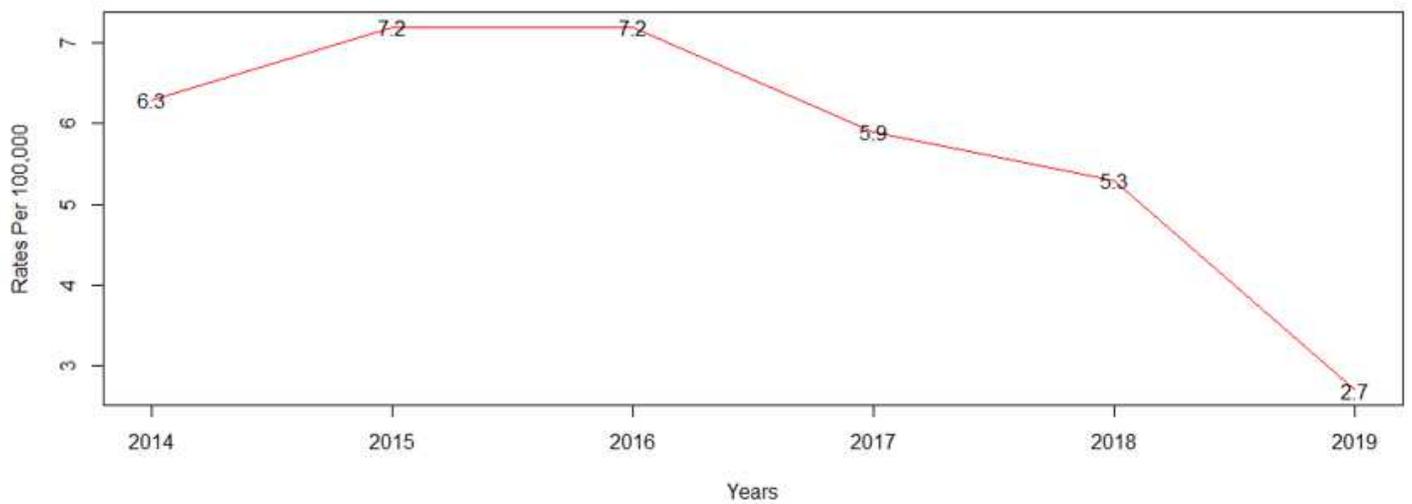
Figure 1

Relative yearly variations in distinct vehicles and their flows



**Figure 2**

Yearly percentages of accident



**Figure 3**

The rates of accidents per 100K

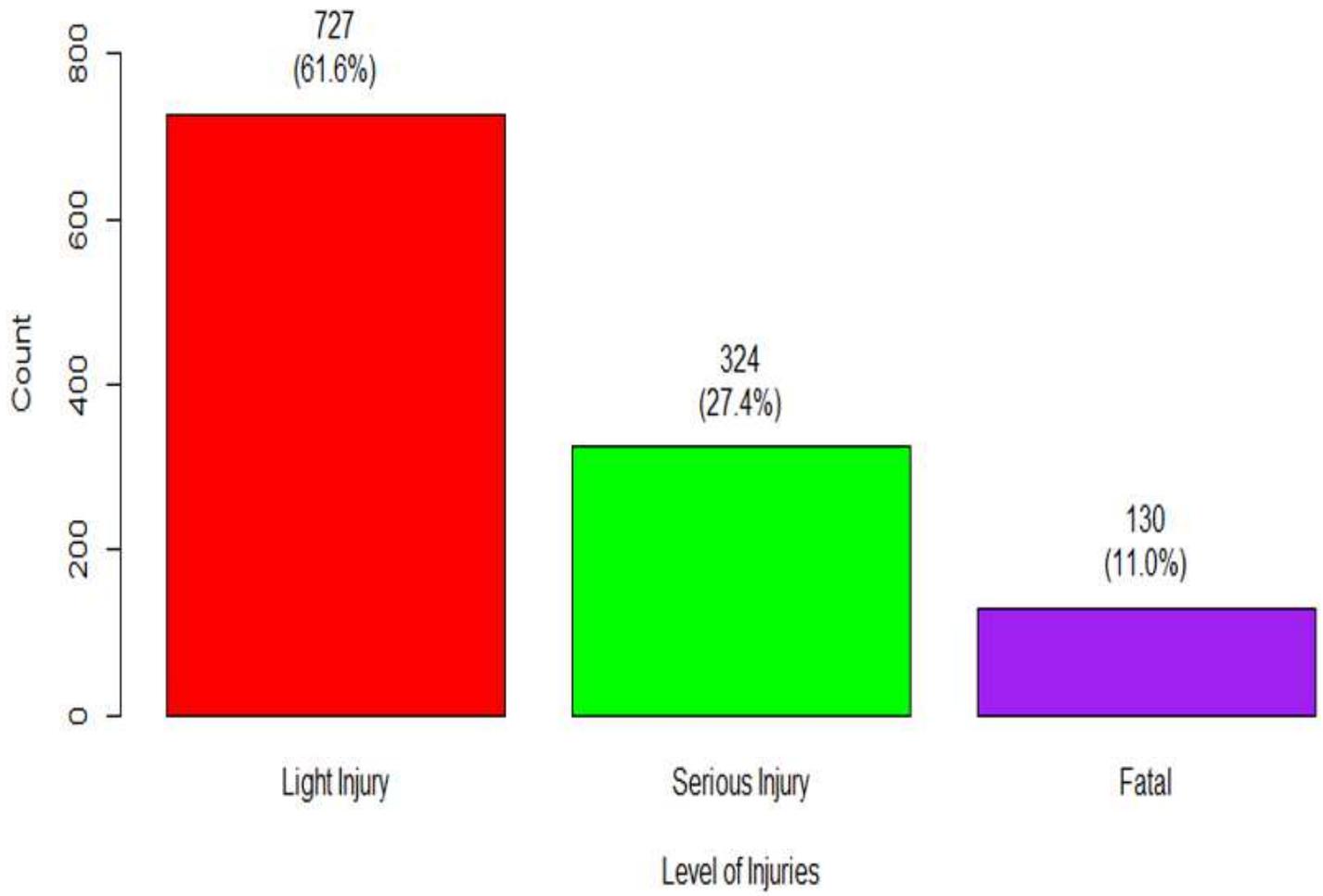


Figure 4

Level of injuries caused from traffic accidents

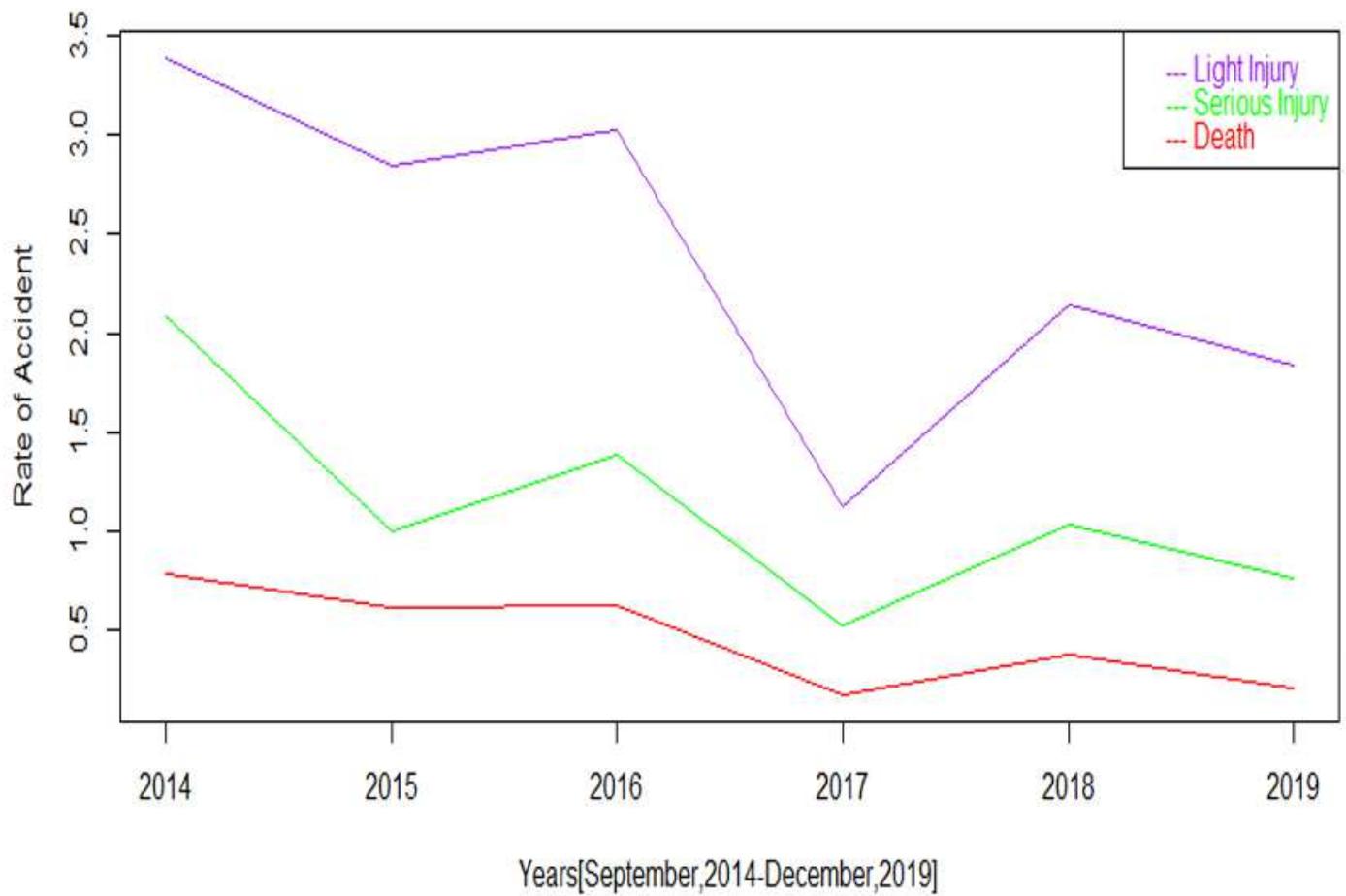


Figure 5

The rates of injury levels

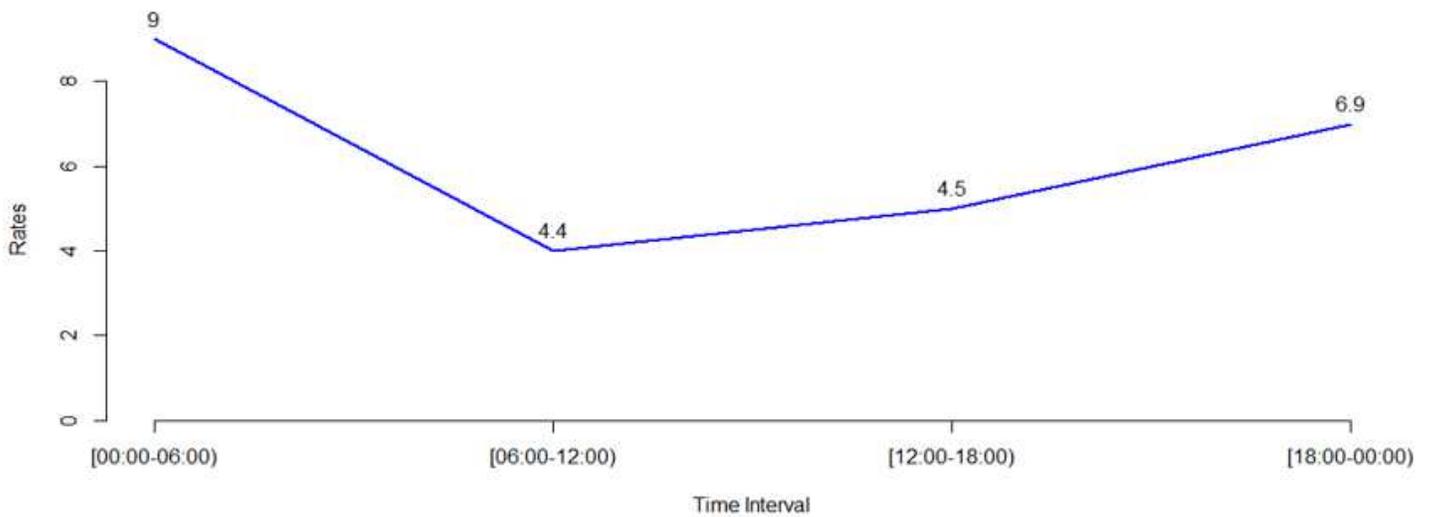


Figure 6

## Time interval rates of accident per 100K flows

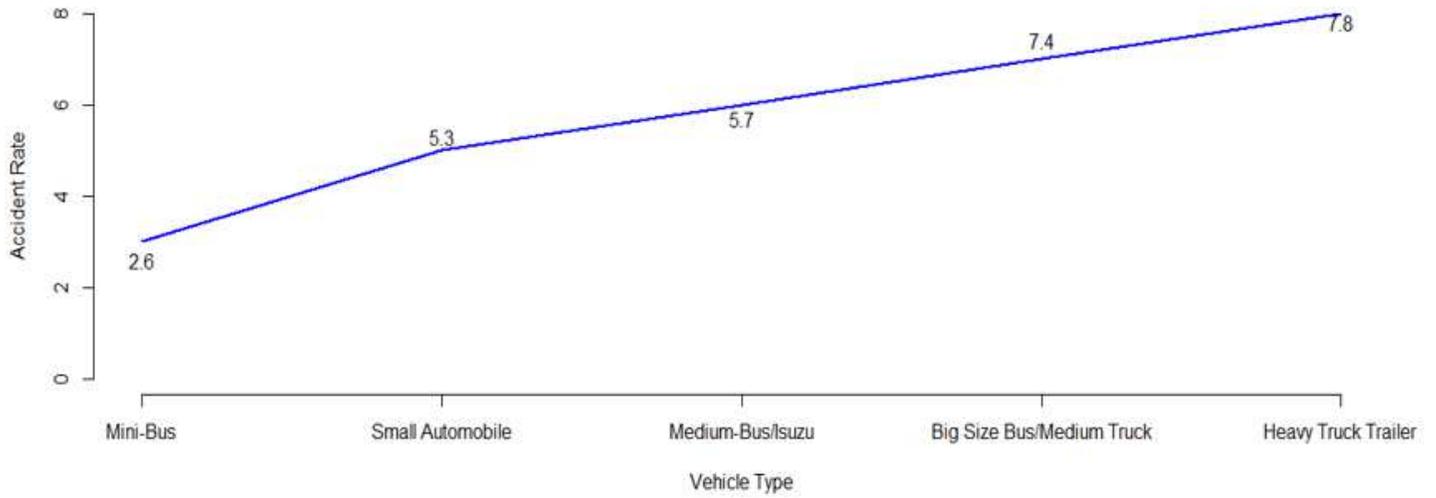


Figure 7

## Vehicle type variation of accidents per 100K flows

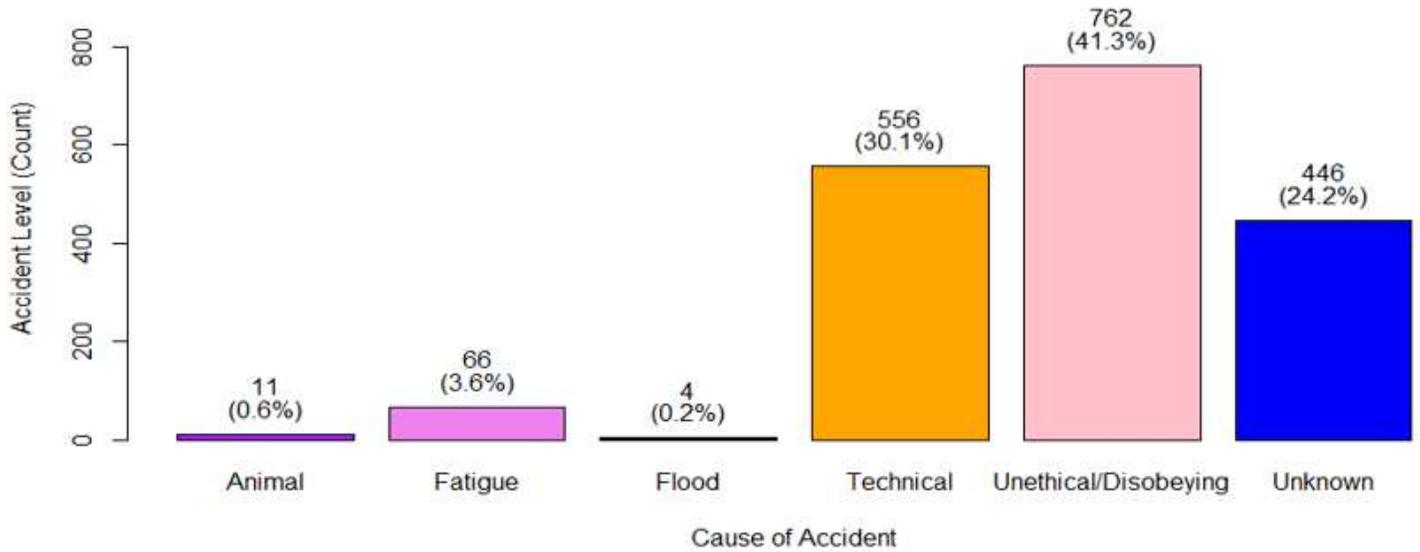


Figure 8

## Cause of accidents

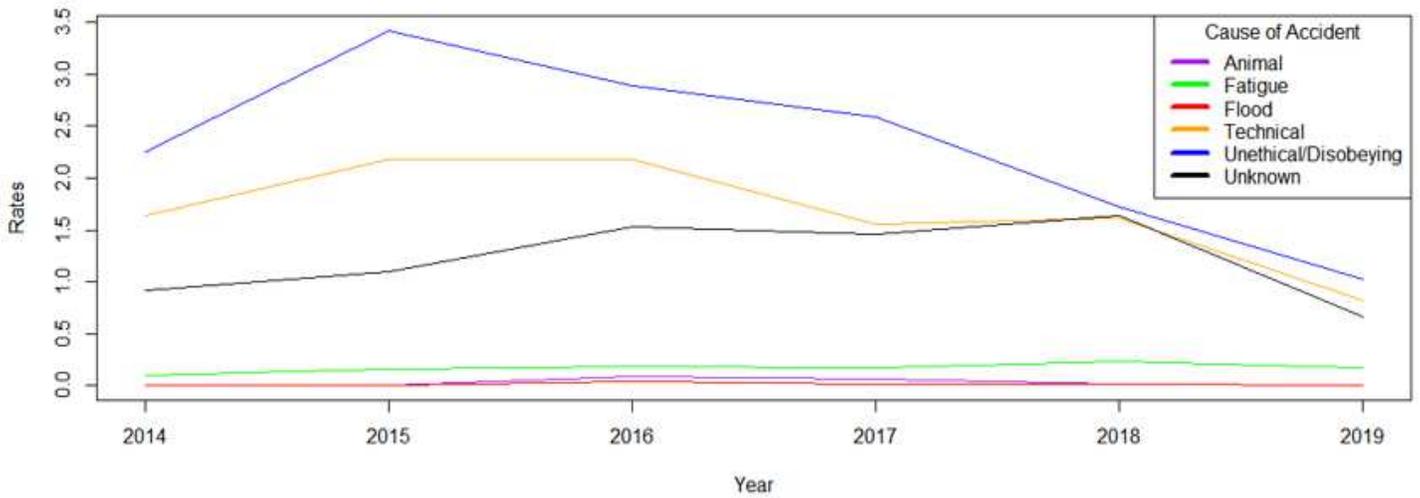


Figure 9

Cause specific accident rates

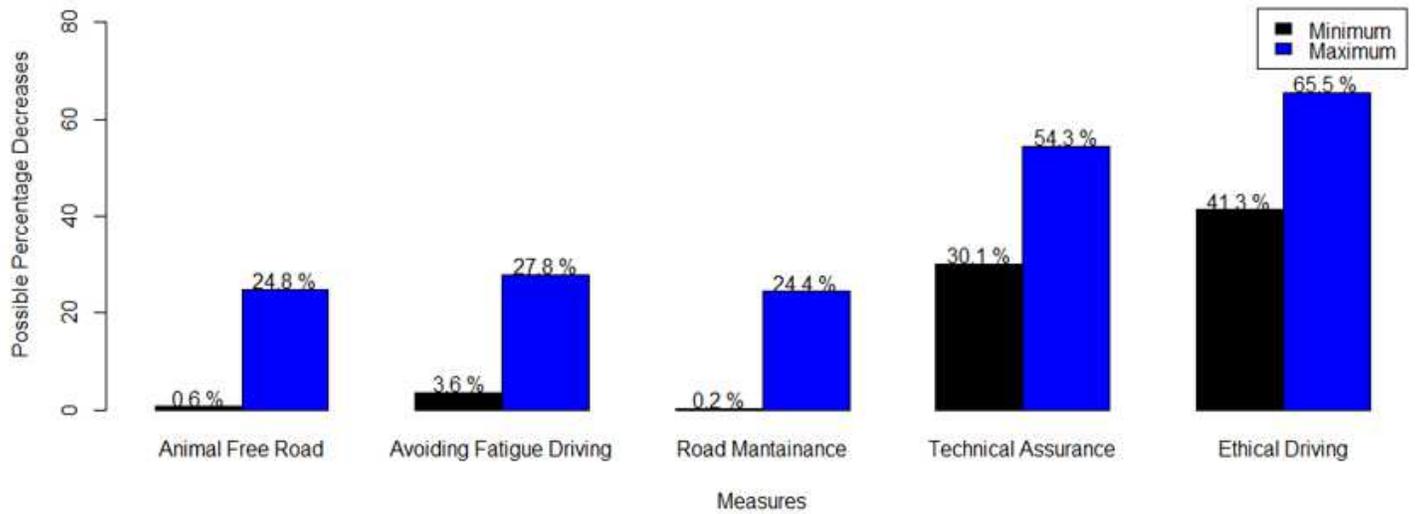


Figure 10

Possible measures to decrease accidents

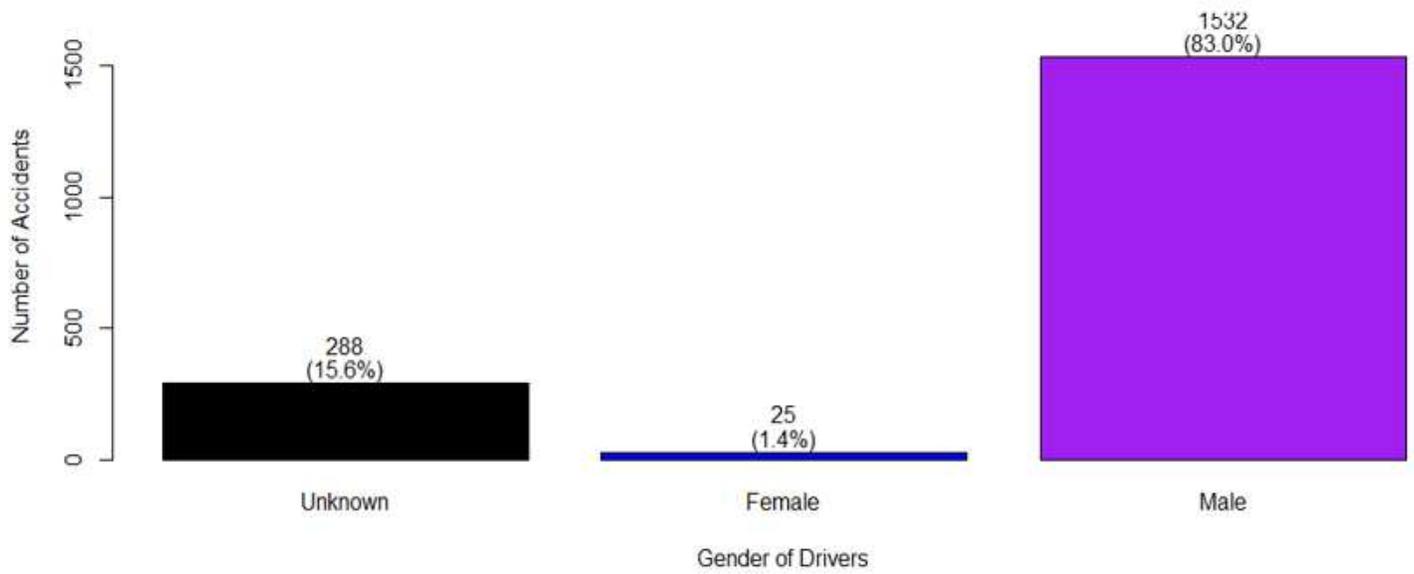


Figure 11

Gender of drivers

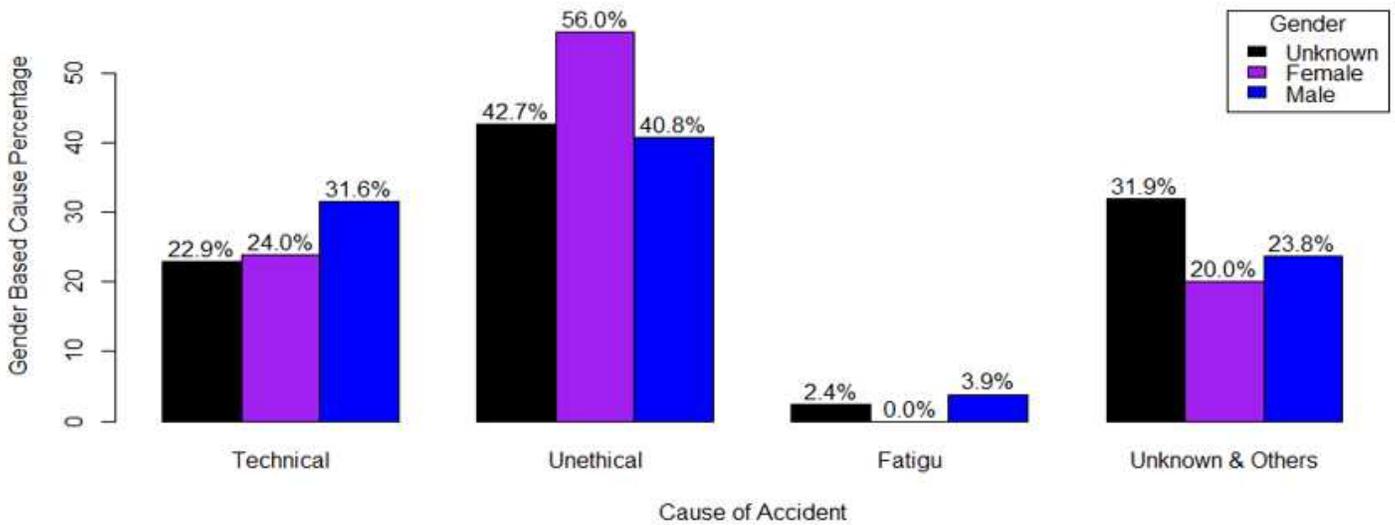
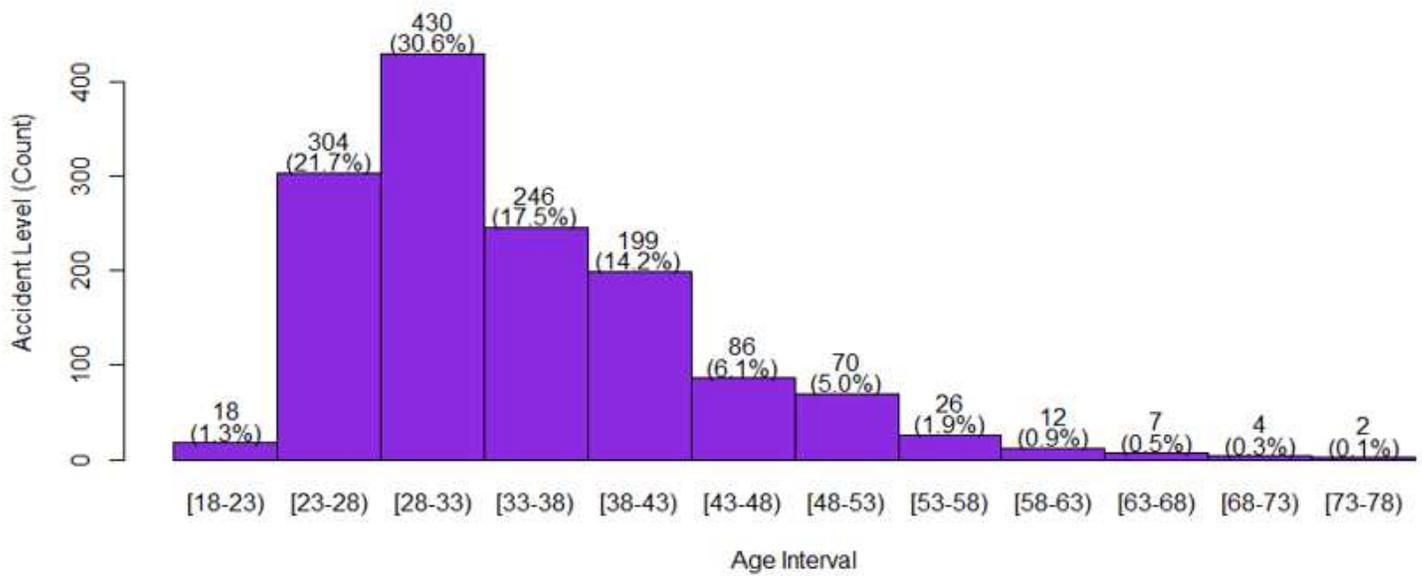


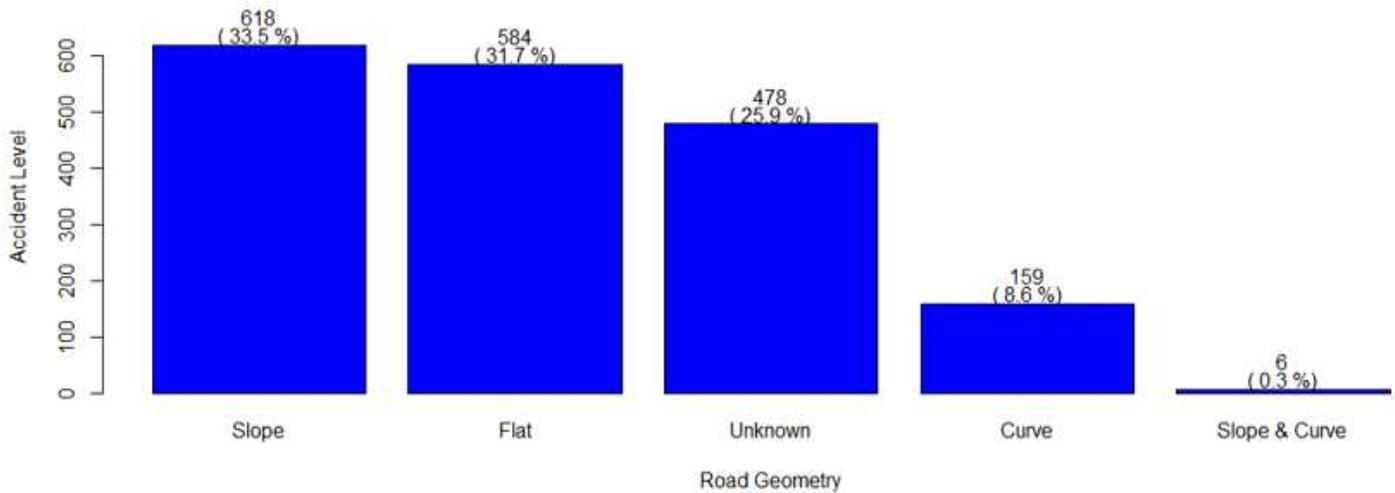
Figure 12

Cross classification of accident causes by gender



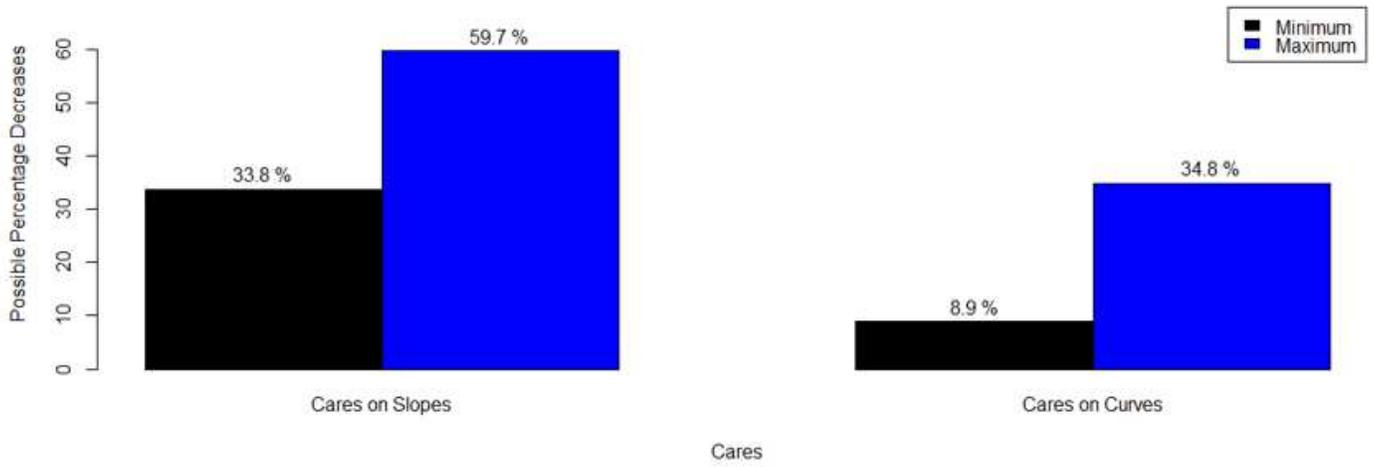
**Figure 13**

Drivers' age distribution



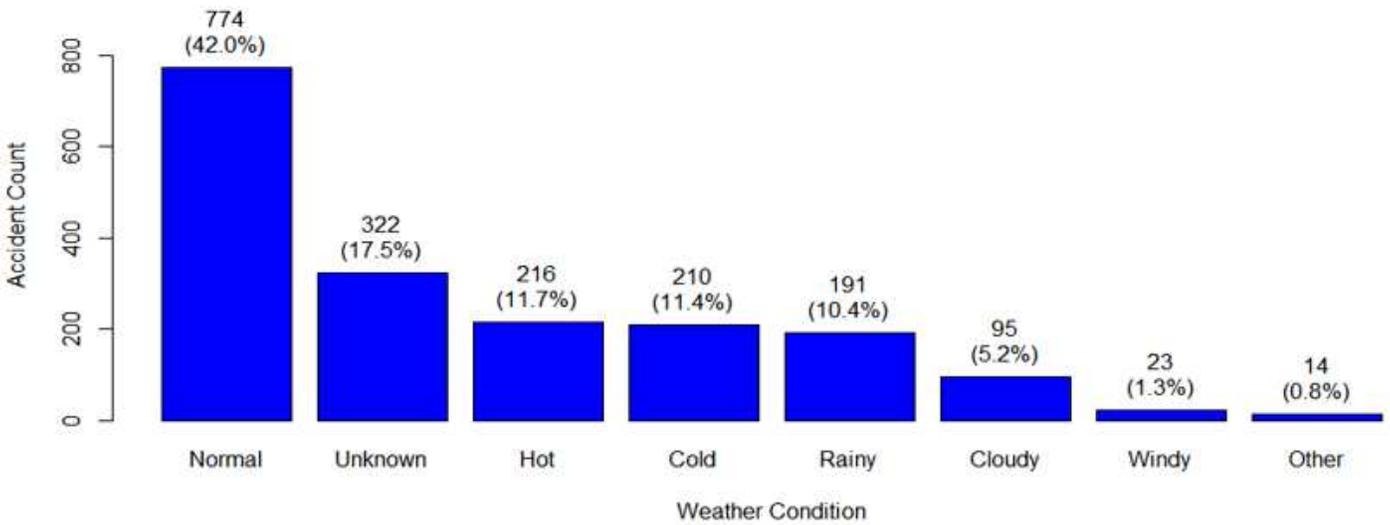
**Figure 14**

Accident road geometry



**Figure 15**

Percentages of accident decreases by taking cares at slopes and curves



**Figure 16**

Accident time weather conditions