A Study on Road Accidents in Kurunegala District in year 2014

D.E.R.S. Nanayakkara
Department of Statistics, Faculty of Science,
University of Colombo

Introduction

- A road accident occurs when a vehicle collides with another vehicle, pedestrian, animal or with a stationary object
- Traffic collisions may result in injury, death, vehicle damage, and property damage
- Worldwide, motor vehicle collisions lead to death and disability as well as financial costs to both society and the individuals involved

Objectives

- To identify the factors affecting accident severity in Kurunegala district

- To make a comparison between the factors affecting accident severity in urban and rural areas of Kurunegala district

Significance of the study

- Many types of road users are daily subjected to a considerable number of road accidents
- The loss incurred due to these various types of accidents is immense
- Limited research is available on the field of road accidents in Sri Lanka
- Due to Kurunegala district's status as a cross roads district, more severe accidents take place daily

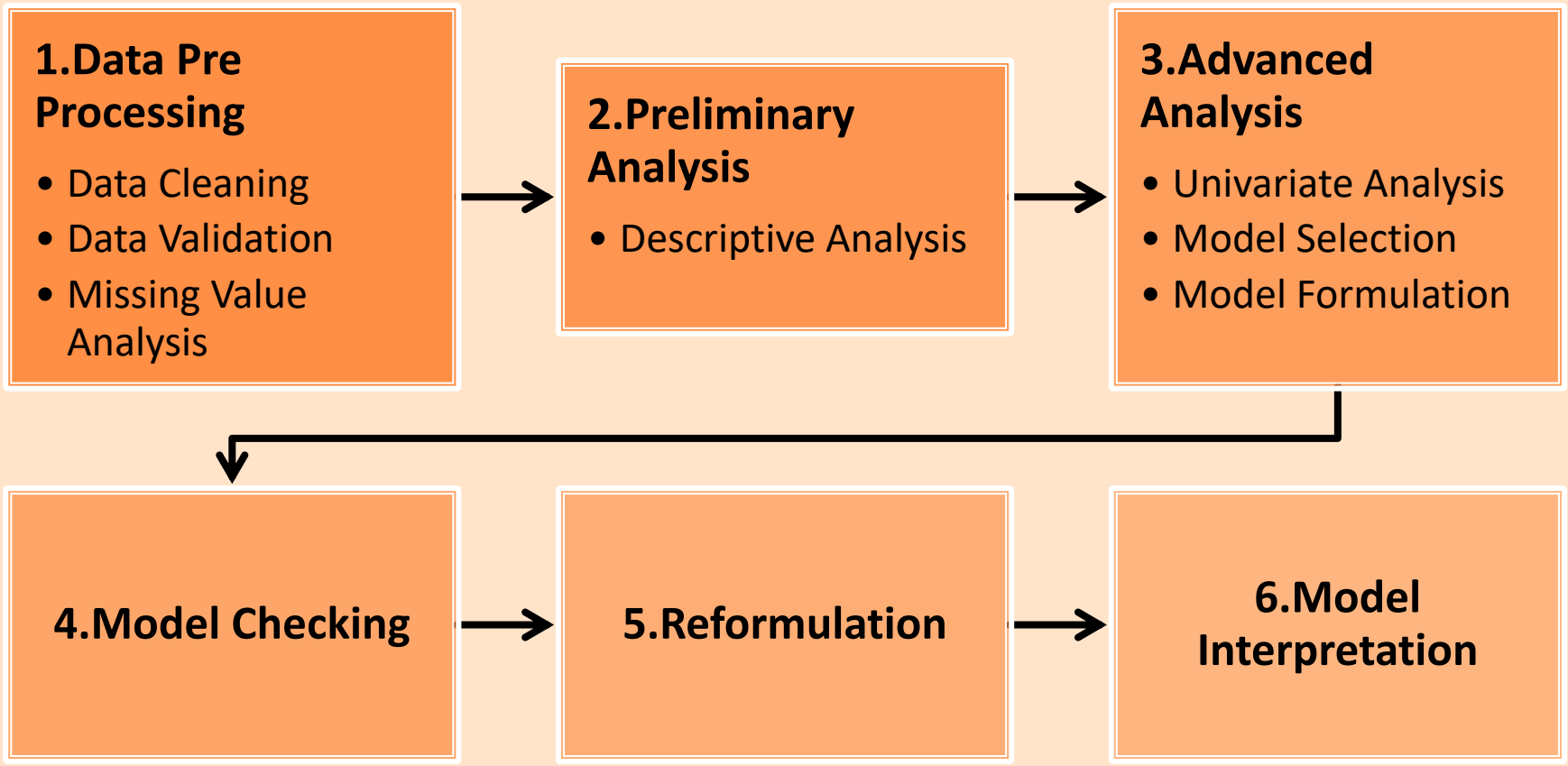
Data

Data Source

- Road accident data for year 2014,collected according to the “297-B” form by the police officers was gained from Kurunegala Traffic Division
- The data base consists of four separate sheets namely,

1. Attendant Circumstances
2. Casualty Details
3. PNT_Accs
4. Vehicle Details

Methodology



Data Pre Processing

OtherCrashI	AlcoholTest	DriverRideratFault	Research Purpose
0	3	1	0
0	3	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0
0	1	1	0
0	1	2	0

Accident Ke	Vehicle Ref	Element Typ	Vehicle Reg	Vehicle Yea	Age of Vehi	VehicleOwr	Direction of	Driver/Pede
14769	1	05	CPMO-4168	0	0	1	S	1
14769	2	19	PTR110	0	0	2	N	1
14770	2	05	MOPED	0	0	1	W	1
14771	1	09	NWNB-0824	0	0	1	W	1
14772	1	05	127-5319	0	0	1	W	1
14772	2	04	0	0	0	1	W	1
14773	1	03	NWGK3162	0	0	1	W	1
14773	2	05	NWBAR1694	2012	2	1	W	2
14774	1	06	SPYI-3210	0	0	1	N	1
14774	2	05	NWBAI7548	2013	1	1	N	1
14775	1	00	0	0	0	0	E	0
14776	1	03	NCGQ-7106	2002	12	1	E	1
14777	1	06	NWQW-7905	2008	6	1	N	1
14778	1	05	NWGH-4732	2001	13	1	S	1
14779	1	03	WPLK-1102	2010	4	1	NE	1
14780	1	09	NWHY-1445	2009	5	1	SW	1

Number of \	HumanPreC	HumanPreC	PedPreCrasl	RoadPreCra	VehiclePreC	CrashFactor	OtherCrashI	AlcoholTest	DriverRider	Research Pu
0	01	00	0	0	0	7	0	3	1	0
0	02	00	0	0	0	0	0	3	1	0
0	02	04	0	0	0	0	2	3	1	0
0	02	03	0	0	0	0	0	3	1	0
1	02	00	0	0	0	7	0	3	1	0
0	03	00	0	0	0	0	0	3	1	0
1	02	00	0	0	1	0	0	3	1	0
6	03	00	0	0	0	7	0	3	1	0
1	02	00	0	0	0	7	0	3	1	0
1	01	02	0	0	1	7	0	3	1	0
0	02	00	0	0	0	0	2	3	1	0
12	02	00	0	1	1	0	1	3	1	0
0	02	00	0	1	1	0	1	3	1	0
0	02	00	0	1	1	0	1	3	1	0
7	02	00	0	1	1	0	1	3	1	0
2	02	00	0	1	1	0	1	3	1	0

Continued…

Accident Key	Number of	Number of	DSDivision
14498	2	1	24
14500	2	1	24
14501	2	2	24
14502	2	1	24
14503	2	0	24
14504	2	1	24
14505	2	1	24
14506	2	1	24
14507	2	1	24
14508	2	2	24
14509	2	0	24
14510	2	1	24
14511	2	1	24
14512	2	1	24
14513	2	0	24
14514	2	2	24
14515	2	0	24
14516	2	0	24
14517	2	0	24
14518	2	2	24
14519	2	1	24
14520	2	1	24
14521	2	0	24
14522	2	1	24

Accident Ke	Vehicle Refi	Element Typ	Vehicle Reg	Vehicle Yea
14498	1	05	SPVQ-1524	2009
14498	2	05	NWUQ-4632	2005
14500	1	06	NWYI-2051	2011
14500	2	02	NWNB-5717	2012
14501	1	06	NWAAI8543	2012
14501	2	05	WPGX-7684	2002
14502	1	05	NWUN-9340	2008
14502	2	04	0	0
14503	1	05	NWAAD3843	2012
14503	2	02	WPNB-2697	2012
14504	1	05	NWGQ-3188	2003
14504	2	04	0	0
14505	1	03	68-4376	1989
14505	2	03	NWHT-7551	2004
14506	1	05	NWXU-8900	2008
14506	2	03	NWPS-2787	2012
14507	1	01	NWKM-0219	2010
14507	2	03	227-1031	1985
14508	1	05	SPWP-7680	2010
14508	2	03	NWGC-0020	2002
14509	1	02	NWPC-4796	2010
14509	2	03	SPLJ-4017	2009
14510	1	06	NWYL-1982	2011
14510	2	09	NWNB-2415	2012

1	Accident_Key	Station_No	Date	Time	Highest_Severity	Urban_or_Rural	WorkDay_or_Holiday	Day_of_Week	Road_Surface	Weather
2	14498	5	1	5	3	2	1	3	1	1
3	14500	1	1	5	3	2	1	4	1	1
4	14501	1	1	5	3	2	1	4	1	2
5	14502	1	1	6	2	2	1	6	1	1
6	14503	1	1	3	4	1	2	1	1	
7	14504	1	1	5	3	2	1	3	1	
8	14505	1	1	5	3	2	1	3	1	
9	14506	1	1	2	3	2	1	4	1	1
10	14507	1	1	5	3	2	1	5	1	1
11	14508	1	1	6	3	1	1	6	1	1
12	14509	1	1	4	4	1	2	7	1	1
13	14510	1	1	5	2	2	2	1	1	1
14	14511	1	1	6	2	1	2	1	1	1
15	14512	1	1	2	3	2	1	2	1	1

Missing Value Analysis

- Quantified extent of missing data
- Characterized types of subjects with missing data
- Founded sets of variables missing on majority of subjects
- List wise deletion was performed for variables whose missing percentages were less than 5%(Acock,2005)
- Unimportant variables whose missing percentages were very high were removed

Missing Value Percentages

Variable	Missing value percentage	Variable	Missing value percentage
Station No	0.00%	Element type	0.52%
Date	0.00%	Age of vehicle	3.32%
Time of accident	0.00%	Vehicle ownership	3.32%
Class of accident	0.00%	Direction of movement	0.70%
Urban/Rural	0.00%	Driver gender	3.14%
Workday/Holiday	0.00%	Driver age	4.80%
Day of week	0.00%	Validity of driving license	4.54%
Second collision	94.6%	Number of years since issue	3.41%
Road surface condition	0.09%	Human pre crash factor 1	30.89%
Weather	0.09%	Human pre crash factor 2	87.43%
Light condition	0.00%	Pedestrian pre crash factor	98.78%
Type of location	0.35%	Road pre crash factor	83.51%
Traffic control	0.17%	Vehicle pre crash factor	82.72%
Posted speed limit signs	0.00%	Crash factor severity	93.98%
Number of casualties	0.00%	Other factors	78.97%
Number of vehicles	0.00%	Alcohol test	0.00%

Continued…

Initial number of observations = 1166

Number of observations after creating the data set = 1146

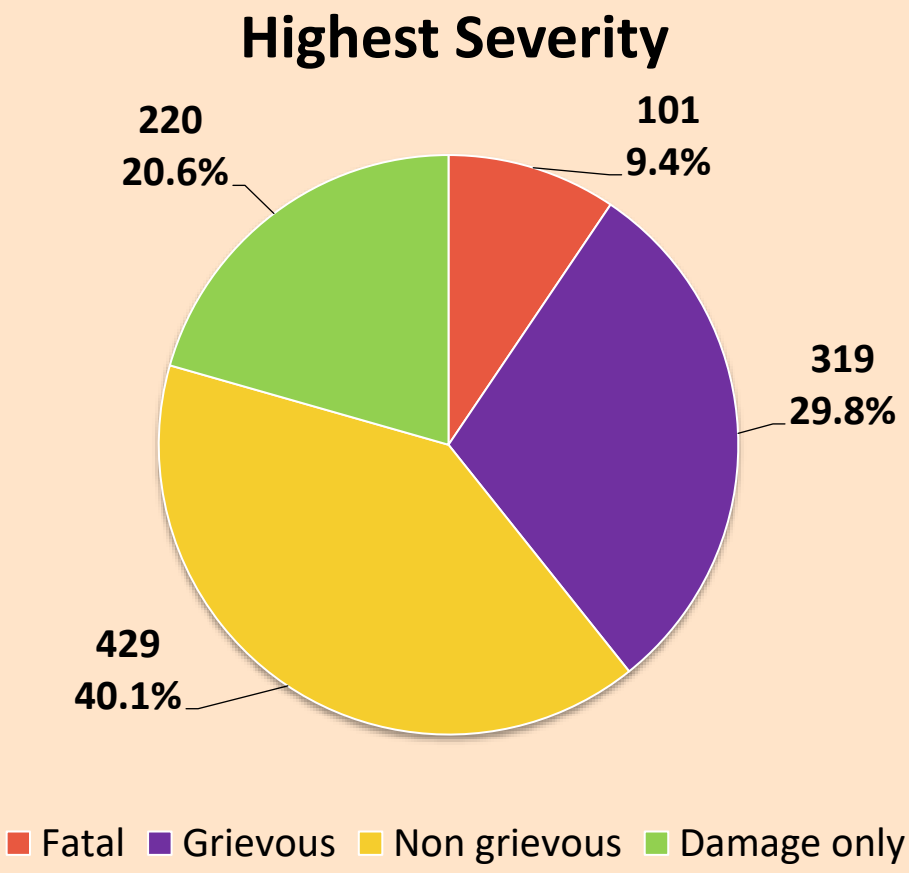
Number of observations after missing value analysis = 1069

Percentage of observations removed due to missing value analysis = 6.72%

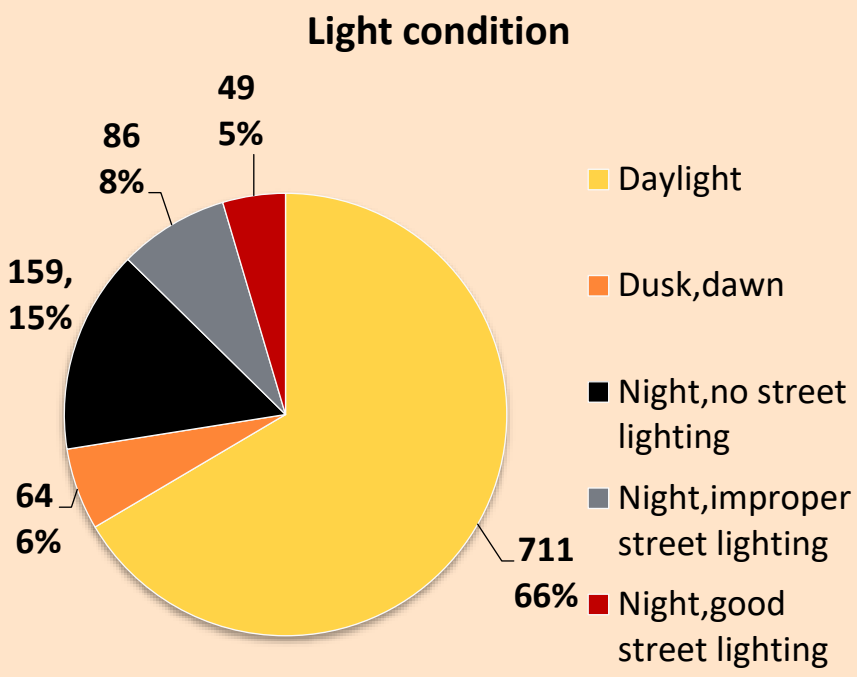
Variables of interest

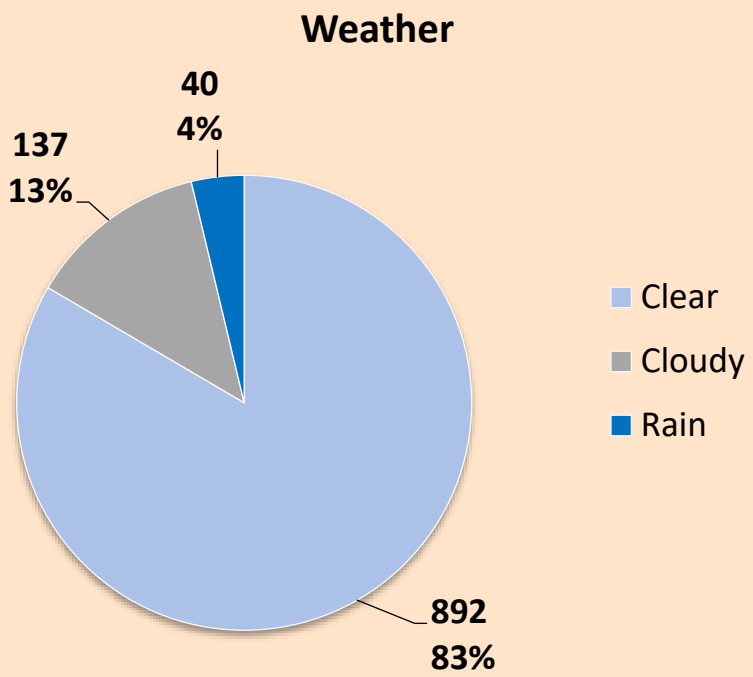
1	Class of accident	Response Variable	12	Traffic control	Traffic Characteristics
2	Station	General Characteristics	13	Posted speed limit signs	
3	Time of accident		14	Number of vehicles	Crash Characteristics
4	Month of accident		15	Number of casualties	
5	Workday/Holiday		16	Element type	Vehicle Characteristics
6	Day of week		17	Driver at fault gender	Driver Characteristics
7	Urban/Rural	Environment Characteristics	18	Driver at fault age	
8	Type of location		19	Validity of driving license	
9	Road surface condition		20	Number of years since issue	
10	Weather		21	Alcohol test	
11	Light condition				

Preliminary Analysis

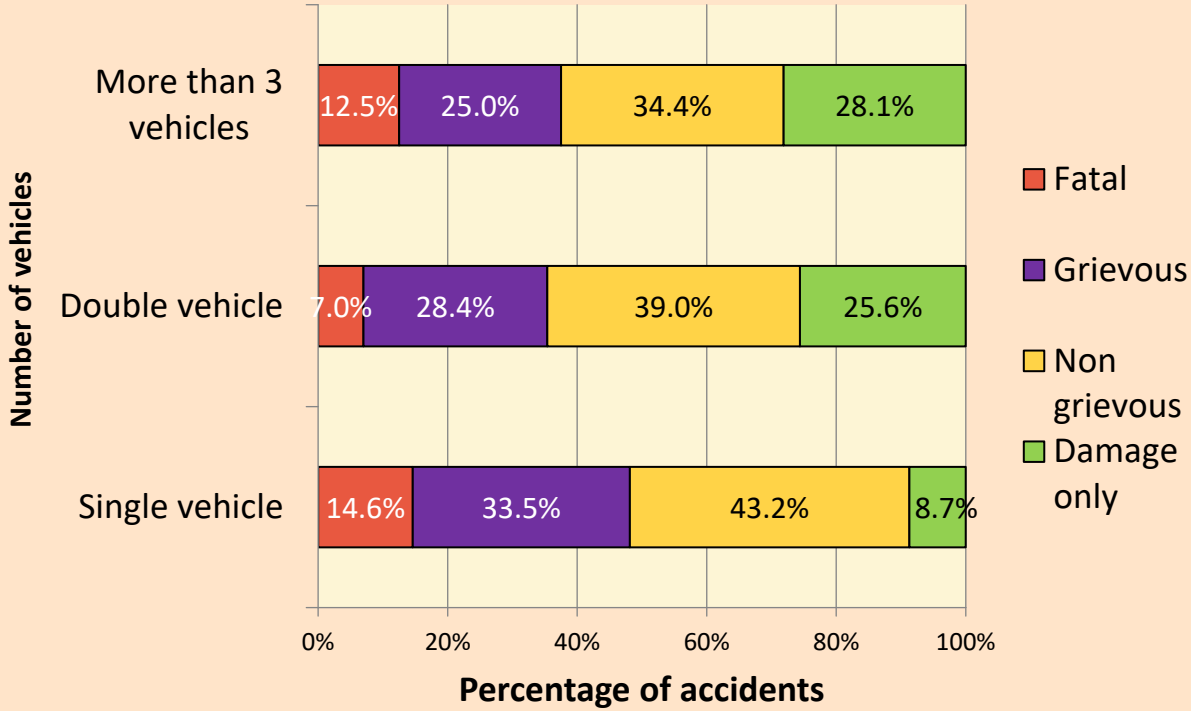


Environment characteristics



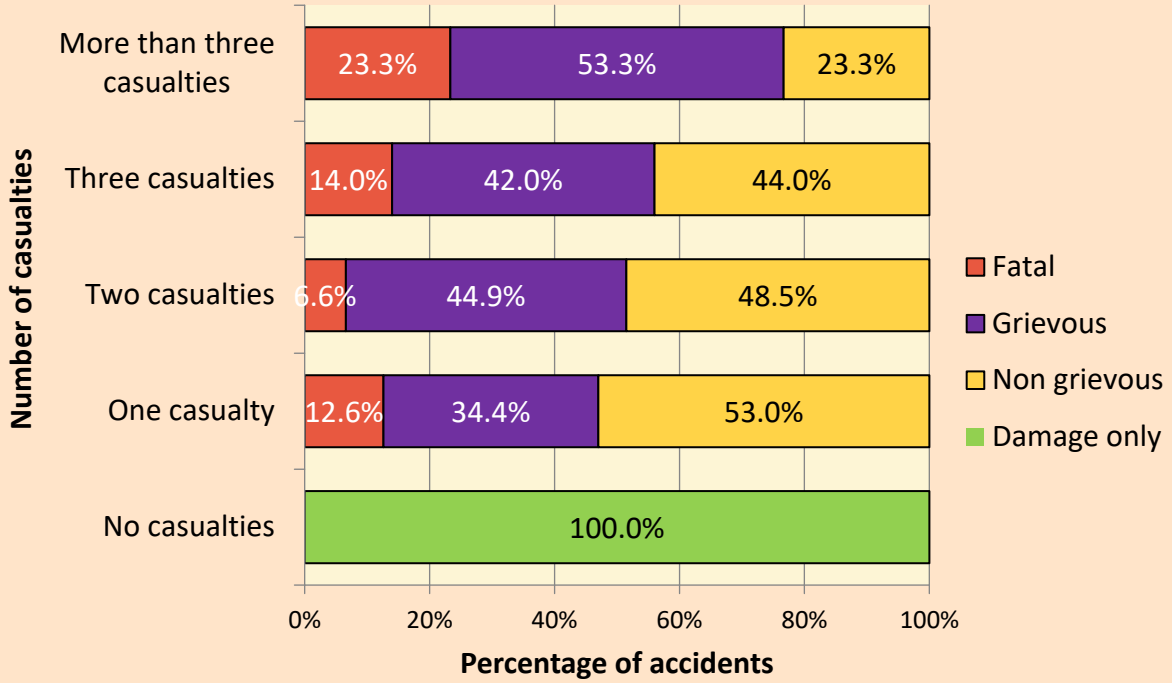


Number of vehicles



- Percentages of occurring fatal, grievous and non grievous accidents are high in single vehicle crashes

Number of casualties



- Percentage of occurring a fatal accident is high when there are more than three casualties
- Percentage of occurring a grievous accident increases with the number of casualties

Advanced Analysis

1. Univariate Analysis
2. Model Formulation
3. Model Validation
4. Reformulation
5. Model Interpretation

Advanced Analysis

Univariate Analysis

Nature of the independent variable	Test
Ordinal with same number of levels as the dependent variable	Kendall's Tau b
Ordinal with different no. of levels as the dependent variable	Kendall's Tau c
Nominal	Kruskal Wallis H
Interval scale	Spearman correlation coefficient

Advanced Analysis

Multicollinearity of the explanatory variables

Ordinal logistic regression model to predict the accident severity

Applying proportional odds model to crash data

Applying continuation ratio model to crash data

Goodness of fit tests for the models

Diagnostic tests

Constructing Receiver Operator Characteristic curve

Advanced Analysis

- Goodness of fit tests
  - Deviance test
  - Pearson test
  - Hosmer and Lemeshow test
- Diagnostic tests
  - Checking the adequacy of the linear predictor
  - Checking for influential observations

Models to predict the accident severity in the whole Kurunegala district

- Model 1:- Fatal vs. Grievous, non grievous and damage only

$$Log\left[\frac{P_{j,k,l,m}}{1-Q_{j,k,l,m}}\right] = \alpha_1 + \beta_j^{Station} + \beta_k^{Traffic\_Control} + \beta_l^{No\_Vehicles} + \beta_m^{No\_Casualties}$$

where j=2,…,10 k=2,3 l=2,3 m=2,3,4

- Model 2:- Grievous vs. Non grievous and damage only

$$Log\left[\frac{P_{a,b,c,d,e,f}}{1-Q_{a,b,c,d,e,f}}\right] = \alpha_2 + \beta_a^{Station} + \beta_b^{Region} + \beta_c^{Weather} + \beta_d^{No\_Vehicles} + \beta_e^{No\_Casualties} + \beta_f^{Element\_Type}$$

where a=2,…,10 b=2 c=2,3 d=2,3 e=2,3,4 f=2,…,10

Models to predict the accident severity in urban and rural areas of Kurunegala district

For urban areas

$$Log\left[\frac{Q_{r,s,t}}{1-Q_{r,s,t}}\right] = \alpha_r + \beta^{Driverfault\_age} + \beta_s^{No\_Vehicles} + \beta_t^{No\_Casualties}$$

Where r=1, 2 s=2, 3 t=2, 3, 4

For rural areas

$$Log\left[\frac{Q_{u,v,w,x,y}}{1-Q_{u,v,w,x,y}}\right] = \alpha_u + \beta_v^{Weather} + \beta_w^{No\_Casualties} + \beta_x^{No\_Vehicles} + \beta_y^{Element\_Type}$$

Where u=1, 2 v=2, 3 w=2, 3, 4 x=2, 3 y=2, 3… 10

Results and Discussion

Odds ratios of the model 1

Odds Ratios Estimates				
Effect	Point Estimate	95% Wald Confidence Limits		$H_0 : \psi = 1$ $H_1 : \psi \neq 1$
		Lower	Upper	
Station 4 vs. 1	13.632*	4.136	44.927	Reject H_0
Station 9 vs. 1	9.324*	2.138	40.667	Reject H_0
Traffic Control 3 vs.1	5.444*	1.662	17.839	Reject H_0
No_Vehicles 2 vs. 1	0.452*	0.291	0.703	Reject H_0
No_Casualties 4 vs. 1	3.293*	1.325	8.185	Reject H_0

Results and Discussion

Odds ratios of the model 2

Odds Ratios Estimates				
Effect	Point Estimate	95% Wald Confidence Limits		$H_0 : \psi = 1$ $H_1 : \psi \neq 1$
		Lower	Upper	
Station 2 vs. 1	3.642*	1.729	7.67	Reject H_0
Station 10 vs. 1	0.363*	0.176	0.747	Reject H_0
Region 2 vs. 1	1.882*	1.273	2.783	Reject H_0
Weather 2 vs. 1	1.863*	1.171	2.964	Reject H_0
No_Vehicles 2 vs. 1	0.584*	0.423	0.806	Reject H_0
No_Vehicles 3 vs. 1	0.321*	0.121	0.849	Reject H_0
No_Casualties 2 vs. 1	2.636*	1.782	3.898	Reject H_0
No_Casualties 3 vs. 1	3.123*	1.589	6.137	Reject H_0
No_Casualties 4 vs. 1	7.315*	2.742	19.518	Reject H_0

Results and Discussion

For urban areas

Odds Ratio Estimates				
Effect	Point Estimate	95% Wald Confidence Limits		$H_0 : \psi = 1$ $H_1 : \psi \neq 1$
		Lower	Upper	
No_Vehicles 2 vs 1	0.382*	0.217	0.674	Reject H_0
No_Vehicles 3 vs 1	0.407	0.098	1.683	Do not reject H_0
No_Casualties 2 vs 1	1.899	0.755	4.78	Do not reject H_0
No_Casualties 3 vs 1	5.861*	1.624	21.156	Reject H_0
No_Casualties 4 vs 1	4.233	0.639	28.047	Do not reject H_0
Driverfault_Age	1.033*	1.011	1.055	Reject H_0

Results and Discussion

For rural areas

Odds Ratios Estimates				
Effect	Point Estimate	95% Wald Confidence Limits		$H_0 : \psi = 1$ $H_1 : \psi \neq 1$
		Lower	Upper	
Weather 2 vs 1	2.073*	1.391	3.089	Reject H_0
No_Casualties 2 vs 1	1.711*	1.181	2.479	Reject H_0
No_Casualties 3 vs 1	2.338*	1.249	4.375	Reject H_0
No_Casualties 4 vs 1	5.24*	2.419	11.352	Reject H_0
No_Vehicles 2 vs 1	0.542*	0.398	0.738	Reject H_0

Conclusions

- The fatal accident percentage of Kurunegala district is higher than the overall fatal accident percentage of Sri Lanka
- The factors, ‘Number of casualties’, ‘Number of vehicles’ and ‘Station’ are associated with the increased accident severity at all levels
- ‘Traffic control’ is associated with the increased accident severity when considered fatal vs. grievous, non grievous and damage only accidents
- The factors ‘Region’, Weather’ and ‘Element type’ are associated with the increased accident severity when considered grievous vs. non grievous and damage only accidents.

Conclusions

- There is a higher likelihood of an accident being fatal when there is no traffic control
- There is a higher possibility of an accident being fatal when it takes place within Wariyapola and Polgahawela police territories with respect to Kurunegala police territory
- The single vehicle accidents and the accidents involving three casualties are more likely to be fatal
- There is a higher possibility of an accident being grievous when it takes place within rural areas as well as in cloudy weather conditions.

Conclusions

Urban areas

- Number of casualties
- Number of vehicles
- Faulty driver's age

Rural areas

- Number of casualties
- Number of vehicles
- Weather
- Element type

Conclusions

- The single vehicle accidents have a higher likelihood of being fatal in both urban and rural areas
- There is a higher likelihood of an accident being fatal in both areas when an accident involves three casualties and the likelihood of being fatal is comparatively higher in rural areas than urban areas
- The young faulty drivers have a higher likelihood of resulting in fatal accidents within urban areas
- There is a higher possibility of an accident being fatal in cloudy weather conditions within rural areas

References

- Abeysekera, W., & Sooriyarachchi, R. (2008). A novel method for testing goodness of fit of a proportional odds model: an application to an AIDS study. *Journal of the National Science Foundation of Sri Lanka, 36*, 2.
- Acock, A. C. (2005). Working With Missing Values. *Journal of Marriage and Family, 67*(4), 1012-1028.
- Albert, A., & Anderson, J. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika, 71*(1), 1-10.
- Dharmaratne, S. D., Jayatilleke, A. U., & Jayatilleke, A. C. (2015). Road traffic crashes, injury and fatality trends in Sri Lanka: 1938–2013. *Bulletin of the World Health Organization*, *93*(9), 640–647. http://doi.org/10.2471/BLT.14.150193
- Quddus, M. A., Wang, C., & Ison, S. G. (2009). Road traffic congestion and crash severity: econometric analysis using ordered response models. *Journal of Transportation Engineering, 136*(5), 424-435.

THANK YOU !!!