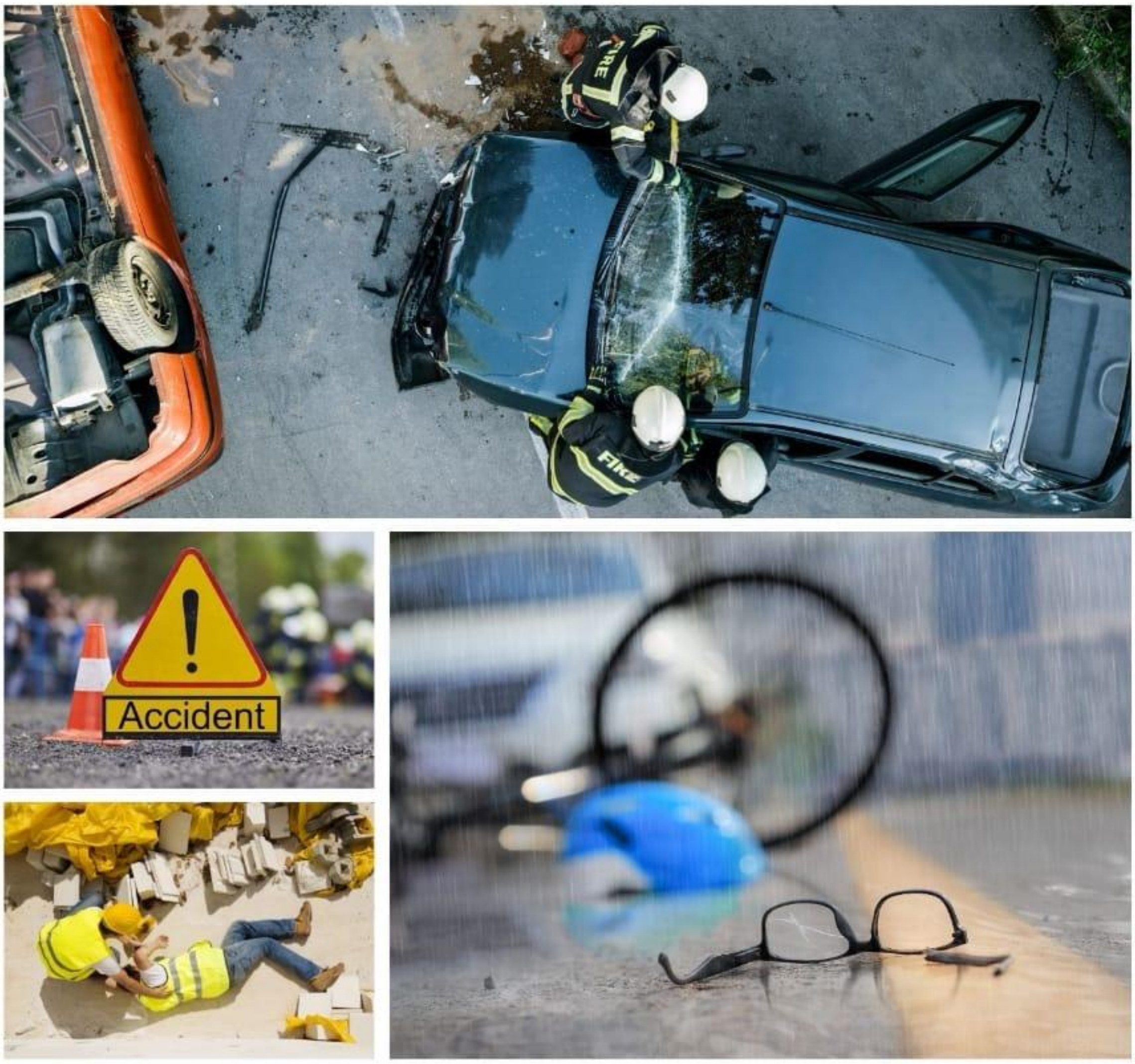


ROAD ACCIDENTS IN INDIA OVER YEARS : A CASE STUDY

PROJECT BY:

Sonam
B.sc(Honours) in Statistics
Registration No: **VB-0995** of **2019-20**
Roll no: B.Sc(Sem-VI)Stat-15.

VISVA-BHARATI UNIVERSITY

ACKNOWLEDGEMENT

I Would like to express my deepest gratitude to my Teacher **SIR SOUMALYA MUKHOPADHYAY** for contributing their valuable time and efforts in helping me out with this project. I would like to express my special thanks to **SIR SUDHANSHU SEKHAR MAITI** and **SIR DEBASISH CHATTERJEE** and all my teachers Their suggestions and feedback have helped me a lot in improving the quality and effectiveness of the project.

I would also like to thank my friends and family for their constant encouragement and support throughout the project.

Lastly, I like to thank all my supporters who have motivated me to fulfil this project before the timeline.

ABSTRACT

Now a days, Road accidents are recognized worldwide as a major problem for people health. Road accidents are one of the main reasons for deaths and disabilities. This imposes the huge socio-economic cost. In India, road accidents are one of the most important causes of death and health loss. This is more among the person of age group 15 to 49 years.

India accounts for about 10% of road accident fatalities of the world every year. This makes India rank 1st in road accident-related deaths. There could be many reasons which lead to the occurrence of accidents in India, some of these reasons are - unfavourable weather conditions, vehicular defects, over speeding, alcohol intoxication of the driver etc. In some cases, the accident could be minor but unfortunately, sometimes it could also be fatal and can lead to major destruction of property. **The main aim of the present study is to suggest the development of a model which deals with the different factors and elements which affects road fatality.** This vital information if given to the concerning authorities, will help them take necessary precautions to prevent the incidence of accidents thereby reducing the fatality rate and destruction caused by road accidents in India.

CONTENTS

INTRODUCTION

DEFINITION OF ACCIDENT:

☛WHO DEFINITION:

Unpremeditated event resulting in recognizable damage.

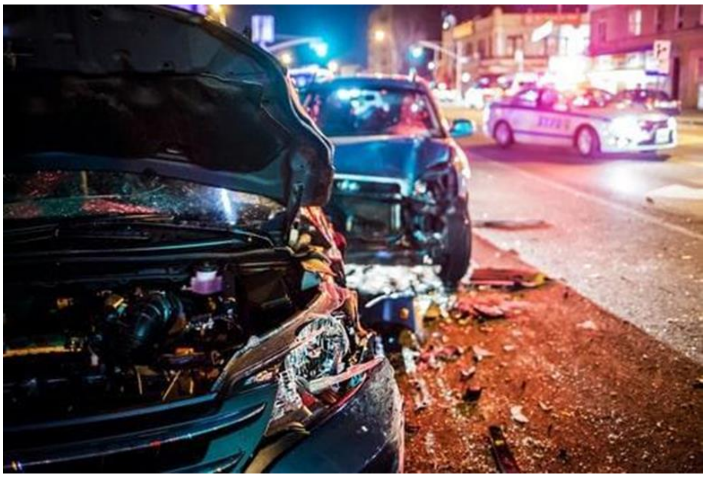
- ☛ **An unexpected occurrence which may involve injury due to collision of vehicle.**

- ☛ **Occurrence in a sequence of events which usually produces unintended injury, death or property damage.**

- ☛ **The RTA full form is “Road Traffic Accident” or “Road and Traffic Authority” or ”Real-Time Access“. A road traffic accident means when a vehicle collided with another vehicle, then it is known as Road Traffic Accident.**

Road Traffic Accidents (RTA) are a global public health concern imposing huge socio-economic costs on a society. Analysis of the incidence of RTA is essential to monitor the effectiveness of different road safety policies implemented to reduce this. The present study analyses RTA mortality reported mainly during 2019 in India.

In India, road transport is the most dominant mode of transportation in terms of traffic share and its contribution to the national economy. With high population growth rates, increasing mobility, and growing numbers of vehicles, tremendous change has occurred to the road transportation network in India over the years. A negative factor associated with the modernization in the road network is the increase in Road Traffic Accidents (RTA). Many people are killed and injured on the roads every day due to RTA leaving behind shattered families and communities imposing huge socioeconomic costs on society. In a developing country like India, roads are the dominant mode of transportation and the road network transports the majority of both passenger and freight traffic**.** The increased amount of motorized traffic on the roads with high acceleration and speed are the critical factors in the causes of RTA injury. The global status report on road safety 2018 by the World Health Organization[1] reported the RTA as the eighth leading cause of deaths globally. The critical and exciting fact about the deaths and injuries due to RTA is that most of them are preventable.



It is necessary to identify the role of the underlying factors that cause the RTA to design and implement policies for reducing RTA. Various factors such as driver's behaviour, vehicle features, road-traffic characteristics and environmental effects act as significant causes for RTA. The use of statistical techniques for analysing RTA data is well established. There have been many studies reported in the literature regarding the modelling of RTA using statistical techniques.

ROAD ACCIDENTS IN INDIA

India's road network, besides being the lifeline of the Nation and a major contributor to socio-economic growth and development, also has the largest contribution to accidental deaths in the country with road accidents accounting for 36-38% (average of 1,50,000 each year) deaths due to other causes during the period from 2015-19.

Road accidents have become a leading cause for fatalities and injuries globally with India being the leading country in this regard. The huge loss of life and attendant economic losses are highly avoidable and require urgent measures to be adopted for effective mitigation

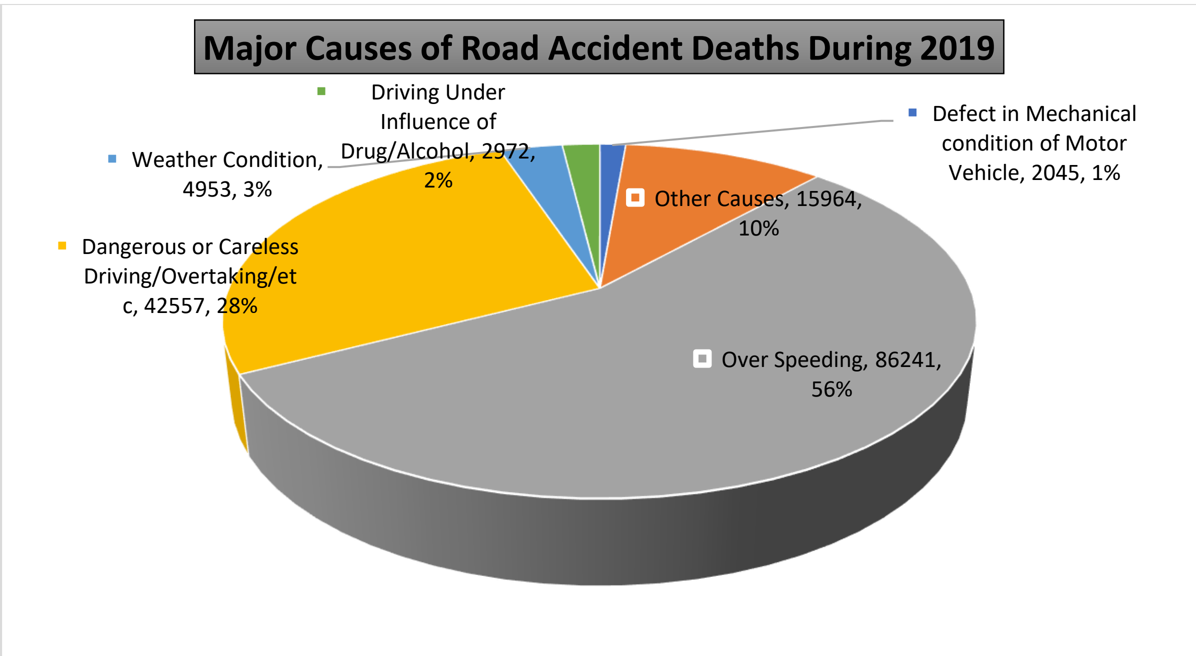
ROAD ACCIDENTS IN 2019.


Img Source: https://bsmedia.business-standard.com/_media/bs/img/article/2014-12/18/full/1418893966-5053.jpg

A total of 4,49,002 road accidents were reported across the country in **2019** resulting in 1,51,113 deaths and 4,51,361 injuries to road users. This averages to almost 1230 accidents 414 deaths daily and 51 accidents and 17 deaths every hour although a marginal decline has been registered when compared to the corresponding figures for **2018.**

CAUSES OF ROAD ACCIDENTS IN INDIA

Road traffic injuries are the eighth leading cause of death globally. In its study highlighting the larger share of Low and Middle Income Countries in road accident fatalities, the World Bank underscores a distinct co-relation between socio-economic status and road use patterns in low and middle income countries such as India. The report states that daily wage workers and those employed as casual labourers in informal activities are more prone to be defined as vulnerable as compared to workers engaged in regular activities. It is often the poor, especially male road-users of working age that constitute the vulnerable road users (VRU) in India where VRUs share road space with other less vulnerable users with their income level having a direct bearing on the mode of transportation used and resultant risk faced by them on that account. Numerous factors can be attributed to be the causative factors of road accidents and can be broadly classified into road environment factors, human factors and vehicular factors



Cause wise analysis of road accidents revealed that most of road accidents were due to overspeeding accounting for 59.6% of total accidents (2,60,898 out of 4,37,443 cases) which caused 86,241 deaths and injuries to 2,71,581 persons. Dangerous/careless driving or overtaking caused 1,12,519 accidents (25.7% of total accidents) which resulted in 42,557

deaths and injuries to 1,06,555 persons during 2019. 2.6% (11,303 out of 4,37,443 cases) of such accidents were due to poor weather condition. Driving under influence of drug/alcohol contributed 1.7% of total accidents which resulted in injuries to 6,675 persons & 2,972 deaths in the country. Cause - wise analysis of fatal road accidents revealed that 55.7% (86,241 out of 1,54,779 deaths) and 27.5% (42,557 out of 1,54,779 deaths) of fatalities in road accidents were due to over-speeding and dangerous/careless driving or overtaking respectively. Poor weather conditions and mechanical defects in motor vehicles caused 3.2% (4,953 deaths) and 1.3% (2,045 deaths) of total deaths due to road accidents respectively during 2019

(in ‘000).

ROAD ACCIDENTS AS PER AGE AND GENDER PROFILE OF FATAL VICTIMS

Majority of the fatal victims belong to the age categories between 18-45 years of age with the numbers remaining more or less the same over 2017 to 2019. 69.3% of the fatalities in 2019 were in the age group of 18-45 years while the working group comprising people of ages 18-60 years formed 84.3% of the total fatalities in the same year. Male road users were among 84-85% of the fatalities on account of road accidents.

ACCIDENTS ON ACCOUNT OF ROAD ENVIRONMENT FACTORS.

ROAD FEATURES. The various road features such as straight stretches, curved roads, location of culverts, bridges, potholes etc also cause accidents. Straight road stretches, which allow for movement of vehicles at high speeds, have accounted for the highest number of accidents (64-66%) in 2018 and 2019. Categories such as curved roads, culverts, potholes and ongoing works have shown an increase in 2019.

Figure 1source:https://www.wallsheaven.com/wall-murals/road-signs.-caution-symbol-under-construction,-work-in-progress-

ONGOING CONSTRUCTION WORKS. Ongoing road and other construction works on or astride the road result in availability of restricted space to the road user. Improper road markings, lack of traffic control etc at such sites further complicates the safety environment around these construction sites.

SPEED BREAKERS.

Presence or absence of speed breakers are amongst a major cause of road accidents. Incorrect location, poor construction and possibility to avoid the speed breaker result in a large number of accidents.

Figure 2source:https://www.newsintervention.com/satire-speed-breaker-country/

Figure 3 source:https://stock.adobe.com/in/images/weather-forecast-concept-background-variety-weather-

WEATHER CONDITIONS. Weather impacts not just the road surface condition but also the visibility of the road user thereby increasing chances of road accidents. Heavy rain, dense fog and hail storms reduce visibility and make the road surface slippery thus posing serious risks to the road users.

POOR LIGHTING.

Lack of lighting on roads is a major cause for accidents. Dim lighting, particularly during night and in adverse weather conditions, impinges on visibility and increases the chances of road accidents.

Figure 4source:https://indianexpress.com/article/cities/delhi/blame-lack-of-streetlights-on-unpaid-bills-says-delhi-firm-5939449/

Figure 5source:https://englishan.com/road-and-traffic-signs/

LACK OF ADEQUATE ROAD SIGNS. Correctly placed road signs are necessary to provide road users with advance warnings wrt road conditions, ongoing works, Traffic Lights Police Controlled Stop Sign Flashing Signal Uncontrolled No of Accidents No of Deaths No of Injuries road features such as turns and sharp bends etc. Absence of such road signs results in the road users being unaware of the requirement of reducing speed or taking additional care in driving.

SIDEWALKS. Constructed for use by pedestrians, poor maintenance and encroachment by residents and street vendors has resulted in the sidewalks not being available for pedestrian use who then move on the roads thereby becoming vulnerable to accidents.

Figure 6source:https://www.ayresassociates.com/well-designed-sidewalks-keep-shoppers-and-others-on-the-move/

NEIGHBOURHOOD ENVIRONMENT. Residential, institutional and market/ commercial areas tend to have higher amounts of traffic congestion and are therefore more prone to road accidents. However, the data received showed larger share of accidents and persons killed in open areas in both 2018 and 2019, perhaps as open areas have lower enforcement presence and may be prone to driving and traffic rule violations

ACCIDENTS ON ACCOUNT OF HUMAN FACTORS

INVALID DRIVING LICENSE: Vehicles driven by untrained and unqualified drivers are a serious traffic hazard and can cause accidents, death and injuries. Though the problem is basically an enforcement issue, it must also be addressed with better facilities and opportunities for training/skilling and evaluation/ testing. While the number of accidents by owner's with valid/ learner's license has shown a decline, the corresponding figure for individuals without a license has shown an

Figure 7 source:legaldesk.com

increase from 8% in 2018 to 9.9% in 2019 despite a decrease from 2017 (10.4%).

Figure 8 source:https://www.deccanherald.com/content/228292/triple-riding-cost-you-thousands.html

TRIPLE RIDING. This is a major factor accounting for the large number of two-wheeler accidents. Besides being illegal, triple riding amounts to contributory negligence as it renders the vehicle unstable and more accident prone

NON-USE OF SAFETY DEVICES

Helmets and Seat Belts. While safety devices such as helmets and seat belts do not cause accidents by themselves, they are instrumental in reducing the number of fatal and grievous injuries. 29.82% (44,666) of the total road accident-related fatalities in 2019 were on account of the driver/ passenger not wearing safety helmets and indicates a callous attitude amongst the population as well as inadequate enforcement measures.

Figure 9source:shutterstock

VIOLATION OF TRAFFIC RULES.

Over-speeding remains the major cause of road accidents in the country with almost 71% (3,19,028) accidents in 2019 occurring due to high speeds and resulting in the death of 1,01,723 (67.3% of total deaths) persons while causing injuries to another 3,26,850 (72.4% of total injuries) individuals. Lane indiscipline was the next major human factor accounting for 5.4% (27,431) of road accidents, 6.1% (9,201) of total deaths and 5.5% (24,628) of total injuries. Balance violations like drunk driving, jumping of traffic signal and use of mobile phones together accounted for 6% of total accidents and 8% of total deaths although these factors have shown an increase in 2019 from the corresponding figures of 2018 highlighting the need for stricter enforcement measures. Traffic accidents, deaths and injuries on account of other causes such as road environment, vehicular condition etc accounted for 17-18% of the total figures with the figures for 2019 showing a substantial reduction as compared to those caused by human factors

DISTRACTED DRIVING. A distracted driver is a motorist who diverts his or her attention from the road, usually to talk on a cell phone, talk to the passengers, send a text message or eat food or even applying makeup. Distracted driving is especially dangerous because, unlike cases of drunk driving which usually occur at night, automobile accidents caused by distracted drivers can occur at any time of the day. Teens and young adults tend to engage in cell phone tasks much more frequently, in riskier situations than adults and therefore are more likely to indulge in distracted driving.

Figure 10source: pinterest

Figure 11source:geotab

NEGLIGENT PARKING. Vehicles are usually parked negligently on the side of the roads without any kind of warning or reflectors thus causing inconvenience for the easy plying of other vehicles. Also, the vehicle which break down are also left unattended / carelessly parked leading to accidents, fatalities and injuries.

ROAD RAGE. Road rage is aggressive behaviour by a motorist towards other road users. This behaviour includes rude gestures, verbal insults, physical threats or dangerous driving methods targeted toward another driver in an effort to intimidate or release frustration. Road rage can lead to assaults and collisions that result in serious physical injuries or even death.

Figure 12source: Bankrate.com

Figure 14source:AITWA

OVERLOADING/ OVERCROWDING OF PASSENGER VEHICLES. Overloaded luggage and passengers beyond the mandated capacity of the vehicle are also reasons that lead to road accidents and cause fatalities and injuries. Such overloading/ overcrowding disturbs the centre of gravity of the vehicle and causes accidents due to loss of balance. Hindrance to the driver’s view of the rear is also a result of overloaded passenger vehicles.

IMPROPER USE OF HEADLIGHTS. High-beam from headlights, particularly used during low visibility, is one of the causes of traffic accidents at night. Though the high-beam lights in cars are prohibited within city limits, people continue to use them even when not required.

Figure 15source:iStock

NOT CROSSING ROADS AT PEDESTRIAN CROSSINGS. The high quantum of fatalities and injuries to pedestrians are on account of non-use of designated crossing places along the roads

ACCIDENTS IN OVER-AGE VEHICLES. Old vehicles are relatively more prone to breakdown and malfunction and therefore require greater care and maintenance on the part of the owner. A study of the data for 2019 reveals that vehicles in the 10-15 year age range accounted for 12.5% of total accidents and 12.6% of total deaths while those above 15 years of age were involved in 11% accidents and 12.3% deaths.

Figure 16source:Dreamstime.com

Most of the causative factors are indicative of the fact that the road accidents in India are caused due to either road environment related issues or due to human factors

PROJECT OBJECTIVE

The objective of my project work are as follows:

1) To check if there is any association between the age and sex of the person died in road fatality for the year 2019.

2) To check whether the rate of accidents in rainy season differ significantly from that in other seasons, for the year 2019

3)To check whether the average rate of change in road accident cases in sunny and rainy season for the interval 2019-2020

ABOUT THE DATA: The data is collected from the website indiastat.com

State-wise Number of Road Accidents Classified by Type of Weather Condition in India - Part I -2019

States/UTs	Sunny/Clear					Rainy				
	Total Accidents	Persons Killed	Persons Injured			Total Accidents	Persons Killed	Persons Injured		
			Greviously Injured	Minor Injury	Total Injured			Greviously Injured	Minor Injury	Total Injured
Andaman and Nicobar Islands	182	17	59	105	164	48	3	15	28	43
Andhra Pradesh	17340	6317	4039	15619	19658	874	241	532	475	1007
Arunachal Pradesh	62	29	41	34	75	49	26	34	29	63
Assam	4322	1735	3287	457	3744	1862	660	1597	176	1773
Bihar	4217	3156	2987	138	3125	2203	1515	1465	91	1556
Chandigarh	299	103	16	253	269	4	0	0	5	5
Chhattisgarh	9033	3209	1540	6726	8266	1535	520	315	1043	1358
Dadra and Nagar Haveli	32	21	58	15	73	17	12	10	0	10
Daman and Diu	52	20	37	13	50	9	0	7	5	12
Delhi	2636	699	400	1977	2377	800	201	96	646	742
Goa	3131	276	245	1091	1336	309	21	20	92	112
Gujarat	15014	6462	8221	5896	14117	1364	656	875	596	1471
Haryana	4619	2347	836	3121	3957	1555	702	475	908	1383
Himachal Pradesh	2501	969	1195	2916	4111	183	64	64	325	389
Jammu and Kashmir	3838	594	2992	1596	4588	148	17	66	119	185
Jharkhand	2363	1602	1404	157	1561	828	699	601	132	733
Karnataka	35753	9721	23394	20170	43564	3181	852	2448	2115	4563
Kerala	35197	3561	28068	11601	39669	4571	650	3513	1455	4968
Lakshadweep	1	0	1	0	1	0	0	0	0	0
Madhya Pradesh	34681	7596	4458	31671	36129	5549	1181	1126	4794	5920
Maharashtra	31536	12311	18351	9066	27417	1186	400	691	343	1034
Manipur	394	99	170	438	608	148	32	51	179	230
Meghalaya	181	89	44	15	59	20	1	6	7	13
Mizoram	44	37	6	27	33	9	7	2	9	11
Nagaland	130	13	30	52	82	91	8	24	50	74
Odisha	6410	3212	3550	2809	6359	1810	723	1080	887	1967
Puducherry	656	87	297	475	772	143	14	84	105	189
Punjab	3416	2485	1378	690	2068	587	332	245	101	346
Rajasthan	21428	9583	6105	14842	20947	355	133	61	273	334
Sikkim	79	26	87	107	194	21	12	10	19	29
Tamil Nadu	48302	8969	3070	54082	57152	2761	474	357	3170	3527
Telangana	18171	5668	2390	16222	18612	66	20	8	54	62
Tripura	534	185	683	5	688	48	26	42	2	44
Uttar Pradesh	17703	9327	7719	4102	11821	7001	3781	3091	1834	4925
Uttarakhand	1087	694	966	219	1185	162	88	129	24	153
West Bengal	4951	2546	4060	745	4805	328	169	292	50	342
India	**330295**	**103765**	**132184**	**207452**	**339636**	**39825**	**14240**	**19432**	**20141**	**39573**

Accidents classified according to type of weather condition during year -2020										
	Sunny/Clear					Rainy				
	Total	Persons	Persons Injured			Total Accidents	Persons	Persons Injured		
States/Uts	Accidents	Killed	Greviously Injured	Minor Injury	Total Injured		Killed	Grevious ly Injured	Minor Injury	Total Injured
Andaman and Nicobar Islands	122	14	48	73	121	19	0	7	17	24
Andhra Pradesh	14836	5325	3158	11897	15055	1376	500	467	956	1423
Arunachal Pradesh	36	19	26	15	41	26	15	28	16	44
Assam	2792	1168	1867	208	2075	1414	494	1052	112	1164
Bihar	4436	3501	3369	102	3471	1945	1476	1655	46	1701
Chandigarh	152	46	14	131	145	1	1	0	0	0
Chhattisgarh	8426	3212	1171	6473	7644	1376	482	193	979	1172
Dadra and Nagar Haveli	42	29	25	46	71	24	12	24	5	29
Daman and Diu					0					0
Delhi	2385	651	106	1965	2071	265	36	17	308	325
Goa	2185	201	181	622	803	190	22	21	56	77
Gujarat	11348	5286	6090	4000	10090	1332	561	693	571	1264
Haryana	3771	1832	786	2146	2932	1300	597	355	767	1122
Himachal Pradesh	2028	787	958	1947	2905	98	49	33	110	143
Jammu and Kashmir	3440	455	2511	1538	4049	161	24	86	118	204
Jharkhand	1828	1353	961	262	1223	737	457	423	104	527
Karnataka	28928	8390	14995	18421	33416	3926	995	2237	2343	4580
Kerala	23945	2353	19279	7096	26375	2799	427	2083	883	2966
Lakshadweep	1	0	0	1	1	0	0	0	0	0
Madhya Pradesh	28497	6883	3483	25271	28754	5323	1420	912	4834	5746
Maharashtra	23945	11061	13301	5720	19021	700	346	460	133	593
Manipur	284	71	145	308	453	73	29	19	77	96
Meghalaya	65	44	35	7	42	11	9	5	5	10
Mizoram	35	19	9	18	27	12	17	10	10	20
Nagaland	203	26	40	96	136	129	11	18	43	61
Odisha	6471	3186	3222	2523	5745	1220	573	707	381	1088
Puducherry	519	75	253	354	607	88	21	21	91	112
Punjab	2422	1888	801	566	1367	561	437	186	166	352
Rajasthan	16579	7989	3805	10670	14475	202	67	45	119	164
Sikkim	66	24	36	46	82	32	11	21	38	59
Tamil Nadu	36926	6336	2523	38686	41209	4673	903	549	4587	5136
Telangana	15734	5742	2103	13309	15412	46	15	5	36	41
Tripura	393	148	377	8	385	29	21	30	1	31
Uttar Pradesh	12860	7266	5225	3169	8394	5633	3021	2463	1415	3878
Uttarakhand	767	517	469	165	634	145	77	99	8	107
West Bengal	4579	2342	3715	475	4190	295	157	270	23	293
India	**261046**	**88239**	**95087**	**158334**	**253421**	**36161**	**13283**	**15194**	**19358**	**34552**

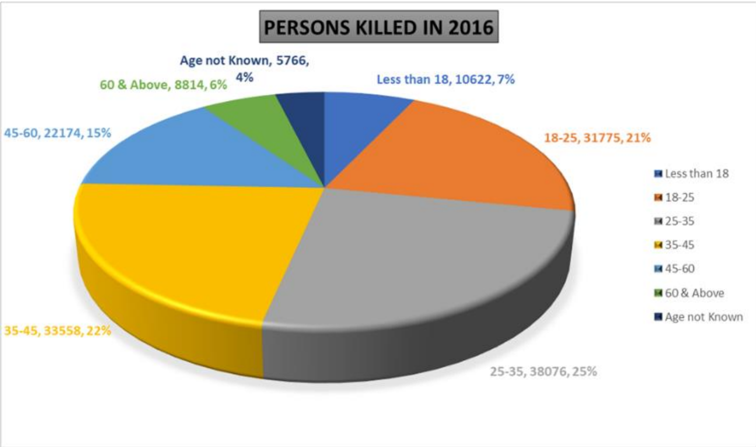
Age Group-wise Number of Accident Victims (Passengers/ Drivers) Killed in Road Accidents in India (2016 to 2019)				
	Persons Killed			
Age Group	**2016**	**2017**	**2018**	**2019**
Less than 18	10622	9408	9977	11168
18-25	31775	34244	32777	33206
25-35	38076	39549	39960	39023
35-45	33558	32788	32672	32509
45-60	22174	22462	22798	22612
60 & Above	8814	9384	9075	9004
Age not Known	5766	78	4158	3591
Total	**150785**	**147913**	**151417**	**151113**

State/Age Group/Sex-wise Number of Drivers Killed in Road Accidents in India -2019														
	Less than 18 Year		18-25 Year		25-35 Year		35-45 Year		45-60 Year		60 and Above		Age Not Known	
States/UTs	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Andaman & Nicobar Islands	0	0	2	0	3	0	3	0	1	0	0	0	0	0
Andhra Pradesh	64	2	699	22	860	28	792	23	546	10	140	6	46	6
Arunachal Pradesh	0	0	17	2	24	2	11	1	6	0	0	0	0	0
Assam	31	2	272	16	326	13	254	9	92	0	2	0	22	0
Bihar	0	0	726	15	605	11	443	10	245	0	35	0	0	0
Chandigarh	2	0	7	0	18	0	9	2	7	1	1	0	0	0
Chhattisgarh	84	0	716	11	829	14	447	7	286	3	56	0	50	3
Dadra & Nagar Haveli	1	0	9	0	3	0	3	0	4	0	0	0	0	0
Daman & Diu	0	0	7	0	1	1	1	0	1	0	0	0	0	0
Delhi	7	1	208	15	221	10	145	7	70	2	9	0	21	0
Goa	2	0	50	2	56	1	30	1	37	0	13	1	0	0
Gujarat	81	5	679	15	891	19	681	7	408	6	82	0	52	1
Haryana	39	5	509	30	473	24	279	10	203	9	100	7	342	18
Himachal Pradesh	16	1	127	1	152	1	101	3	85	0	19	0	0	0
Jammu & Kashmir	16	0	112	2	131	3	65	0	20	0	6	0	5	0
Jharkhand	198	0	357	16	357	15	233	1	176	1	162	2	166	5
Karnataka	68	4	1095	10	1438	12	1072	7	645	5	143	3	0	0
Kerala	30	0	476	5	440	10	313	13	467	14	243	5	28	1
Lakshadweep	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Madhya Pradesh	186	25	1086	33	1320	35	955	18	445	17	143	6	98	3
Maharashtra	62	1	1335	36	1799	55	1518	32	855	11	152	2	90	1
Manipur	0	0	9	1	25	1	19	1	10	0	3	0	0	0
Meghalaya	14	11	12	15	6	7	1	0	4	5	0	0	0	0
Mizoram	3	0	3	1	6	0	5	0	4	0	0	0	1	0
Nagaland	0	0	12	0	2	0	0	0	3	0	0	0	0	0
Odisha	84	2	604	23	643	10	518	14	269	3	36	1	24	0
Puducherry	0	1	14	0	13	2	15	0	10	0	7	0	0	0
Punjab	425	80	579	75	356	112	75	65	74	20	25	27	0	0
Rajasthan	98	0	1340	12	1514	21	942	11	341	5	25	1	10	1
Sikkim	0	0	5	0	10	2	6	0	2	0	2	0	0	0
Tamil Nadu	246	12	613	69	763	62	1191	56	594	31	197	11	0	0
Telangana	54	1	646	6	827	9	616	8	491	7	157	1	23	0
Tripura	1	1	35	1	25	0	8	0	5	0	3	0	0	0
Uttar Pradesh	838	92	2349	190	2301	160	1809	89	1028	66	287	17	193	41
Uttarakhand	1	1	70	3	97	6	81	4	43	1	6	2	2	0
West Bengal	48	3	14	5	507	4	290	4	273	1	38	0	46	0
India	**2699**	**250**	**14794**	**632**	**17042**	**650**	**12931**	**403**	**7750**	**218**	**2092**	**92**	**1219**	**80**

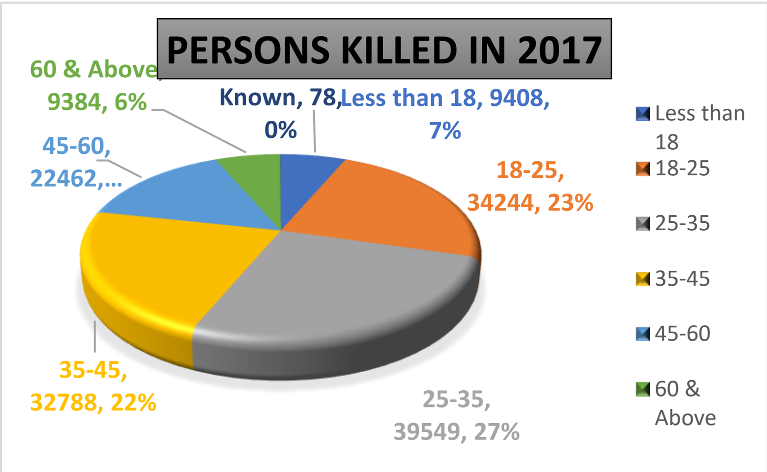
Years	Total Number of Road Accidents (in numbers)	Total Number of Persons Killed (in numbers)	Total Number of Persons Injured (in numbers)	Population of India (in thousands)	Total Number of Registered Motor Vehicles (in thousands)	Road Length (in kms)
1970	114100	14500	70100	539000	1401	1188728
1980	153200	24000	109100	673000	4521	1491873
1990	282600	54100	244100	835000	19152	1983867
1994	325864	64463	311500	904000	27660	2890950
1995	351999	70781	323200	924359	30295	2975035
1996	371204	74665	369502	941579	33786	3202515
1997	373671	76977	378361	959792	37332	3298788
1998	385018	79919	390674	978081	41368	3228356
1999	386456	81966	375051	996130	44875	3296650
2000	391449	78911	399265	1014825	48857	3316078
2001	405637	80888	405216	1028610	54991	3373520
2002	407497	84674	408711	1045547	58924	3426603
2003	406726	85998	435122	1062388	67007	3528654
2004	429910	92618	464521	1079117	72718	3621507
2005	439255	94968	465282	1095722	81502	3809156
2006	460920	105749	496481	1112186	89618	3880651
2007	479216	114444	513340	1128521	96707	4016401
2008	484704	119860	523193	1144734	105353	4109592
2009	486384	125660	515458	1160813	114951	4471510
2010	499628	134513	527512	1176742	127746	4582439
2011	497686	142485	511394	1210193	141865.6	4676838
2012	490383	138258	509667	1208116	159490.6	4865394
2013	486476	137572	494893	1223581	181508	5231922
2014	489400	139671	493474	1238887	190704	5402486
2015	501423	146133	500279	1254019	210023	5472144
2016	480652	150785	494624	1268961	230031	5603293
2017	464910	147913	470975	1283601	NA	NA

Zones	States
Eastern	Bihar
	Jharkhand
	Odisha
	West Bengal
Western	Dadra and Nagar Haveli
	Daman and Diu
	Goa
	Gujarat
	Maharashtra
Northern	Chandigarh
	Delhi
	Haryana
	Himachal Pradesh
	Jammu and Kashmir
	Punjab
	Rajasthan
Southern	Andaman and Nicobar Island
	Andhra Pradesh
	Karnataka
	Kerala
	Lakshadweep
	Puducherry
	Tamil Nadu
Central	Chhattisgarh
	Madhya Pradesh
	Uttar Pradesh
	Uttarakhand
North Eastern	Arunachal Pradesh
	Assam
	Manipur
	Meghalaya
	Mizoram
	Nagaland
	Sikkim Tripura

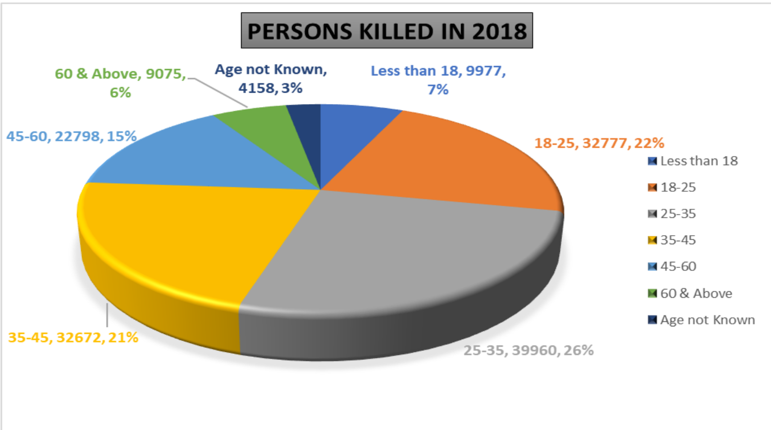
GRAPHICAL REPRESENTATION OF THE DATA



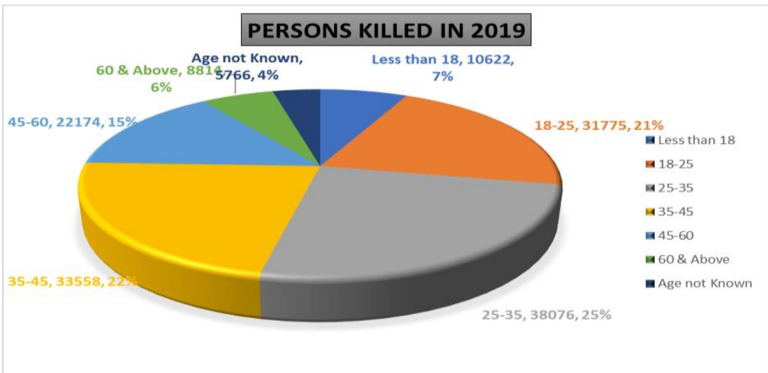
A plot showing the different age groups contribution in road accident fatality in 2016.



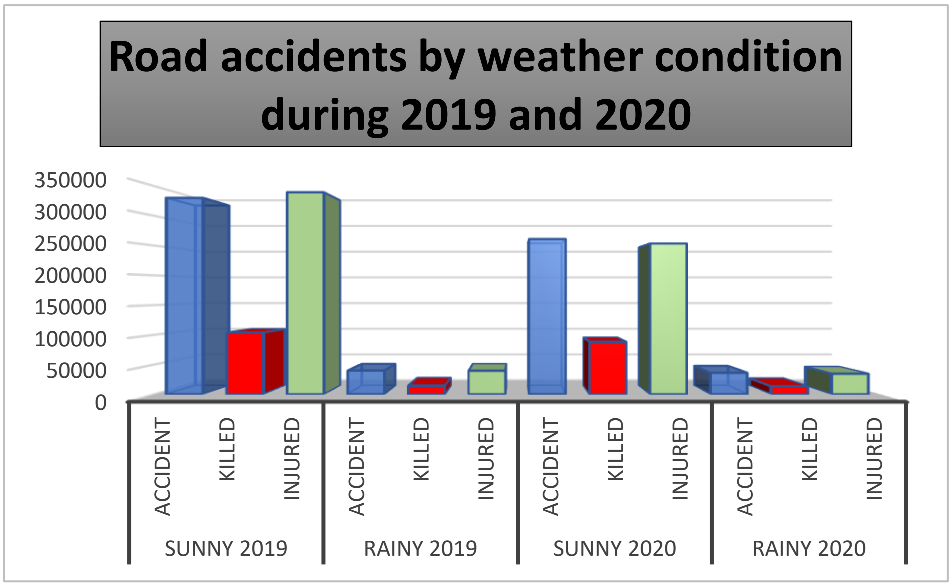
A plot showing the different age groups contribution in road accident fatality in 2017.



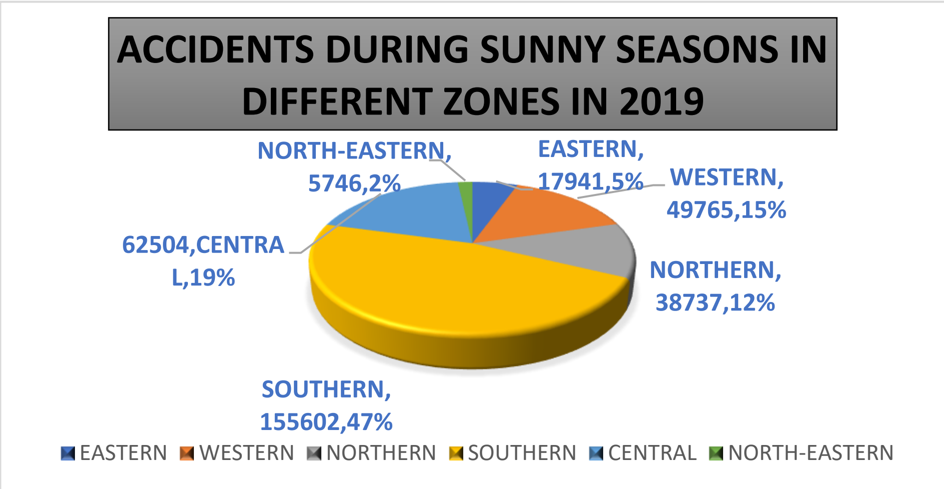
A plot showing the different age groups contribution in road accident fatality in 2018.



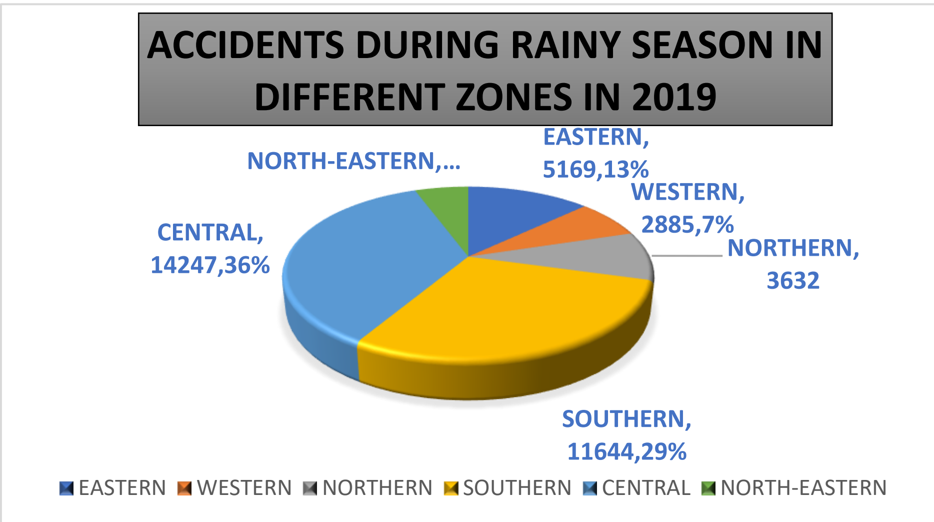
A plot showing the different age groups contribution in road accident fatality in 2019.



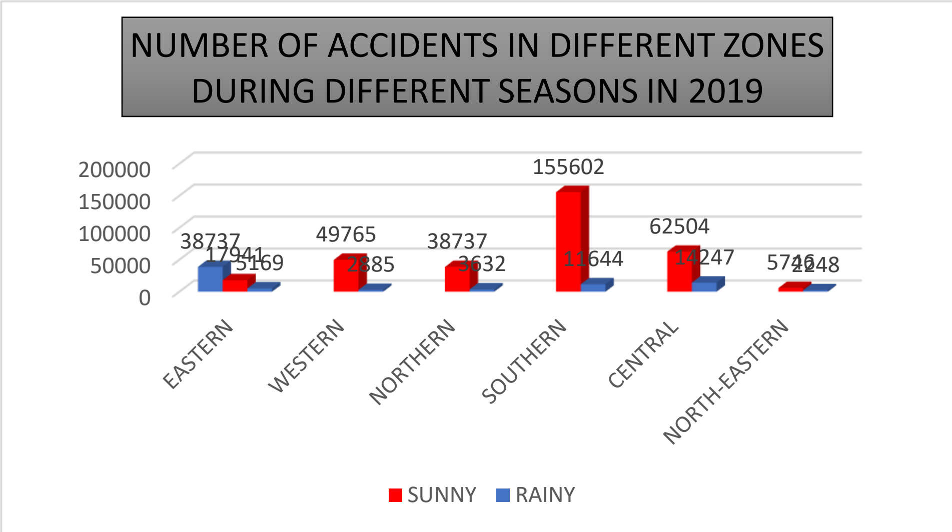
We can see that there is more number of accident cases in sunny seasons than that in rainy season but the cases follow approximately the same pattern i.e.,the average change in number of cases 2019-2020 sunny season and rainy season are the same .



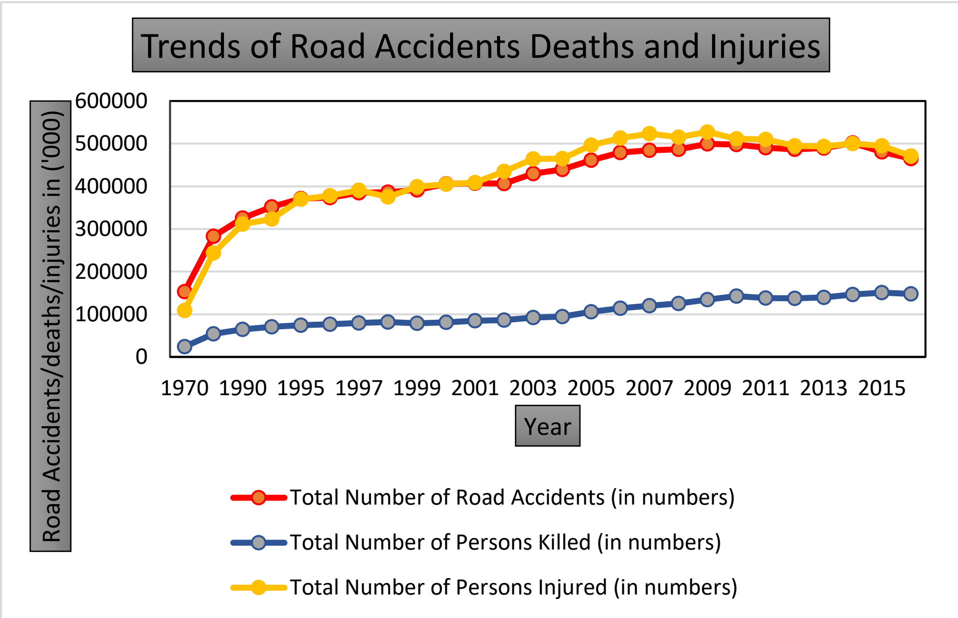
A plot showing the proportions of road accidents in INDIA due to different zones in sunny season in 2019



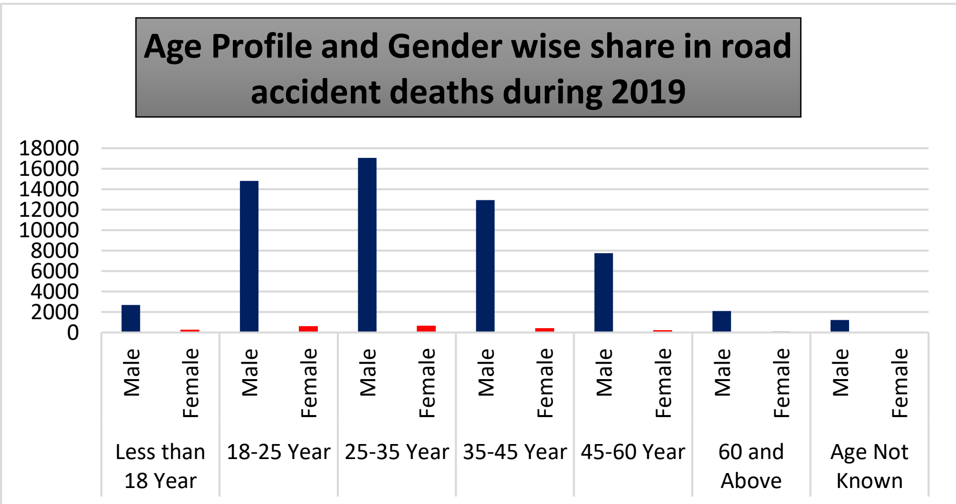
A plot showing the proportions of road accidents in INDIA due to different zones in rainy season in 2019



A plot showing the proportions of road accidents in INDIA due to different zones in both sunny and rainy season in 2019

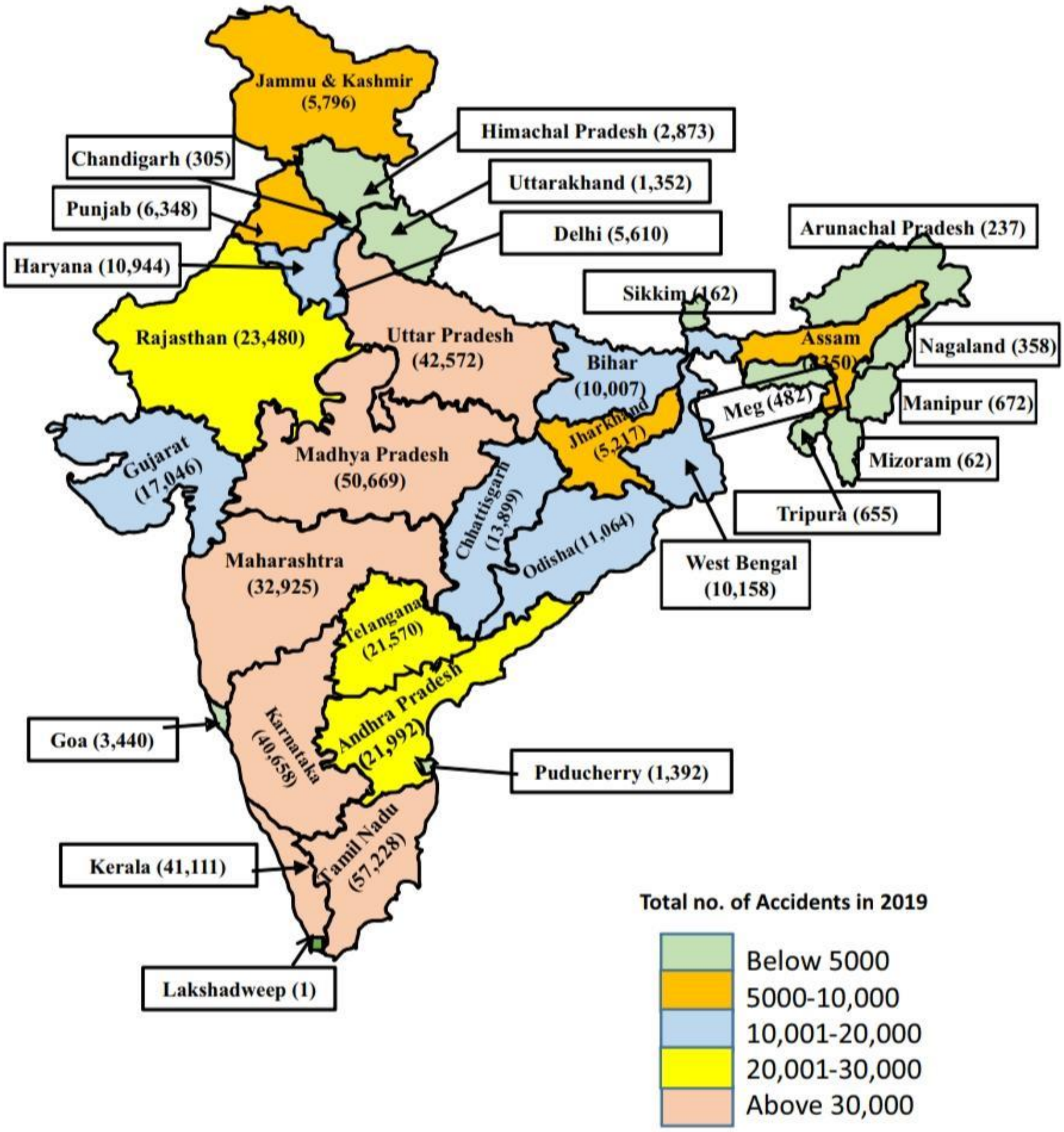


The above graph reveals a consistent increase in road accidents, accident related deaths and road injuries up to **2010** after which all three categories of accidents, deaths and injuries have stabilized with marginal fluctuations and slight decrease in **2017.**

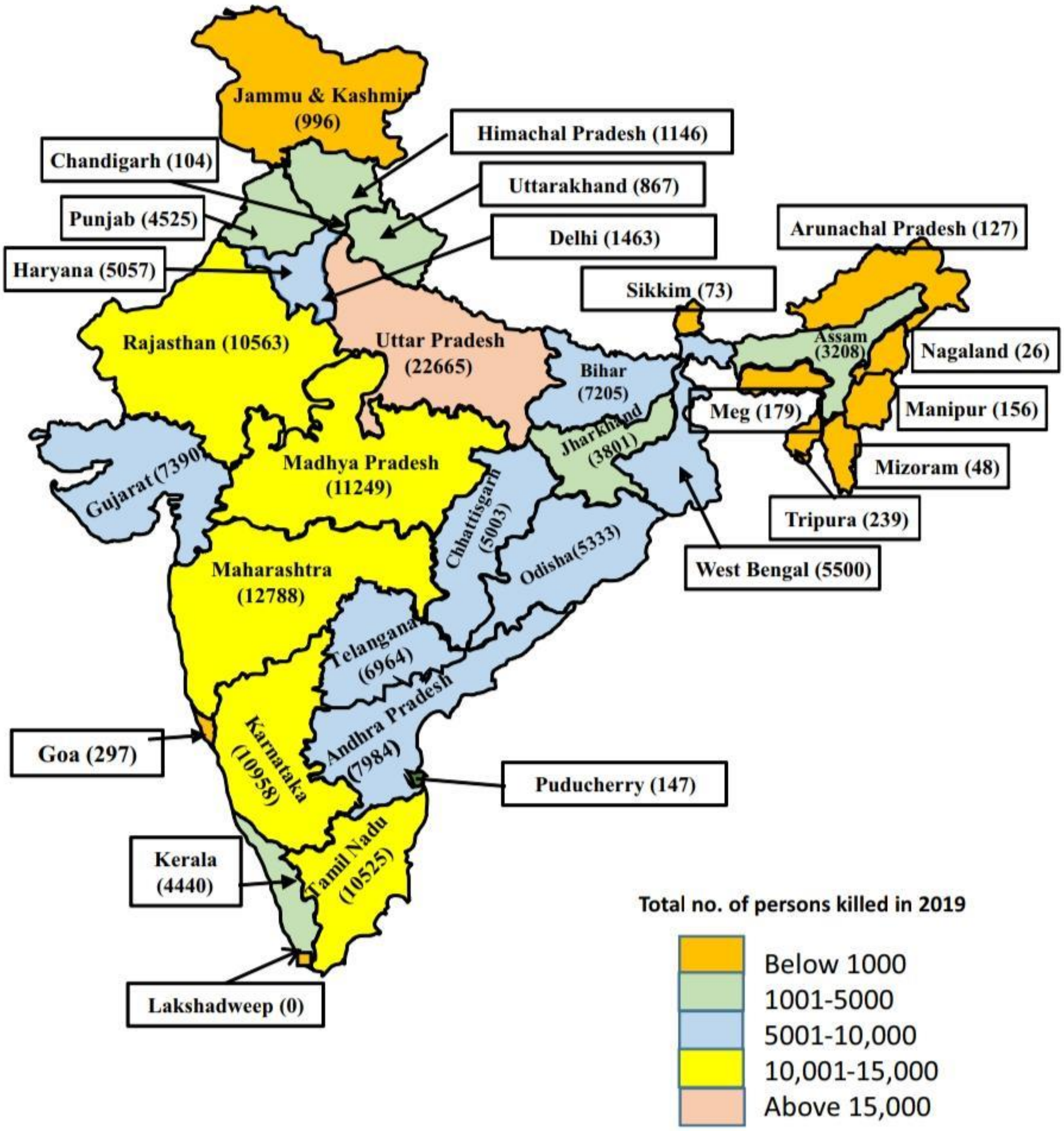


plot showing the proportions of road accidents in INDIA due to different age groups and different gender in 2019

Map 2.1 Road Accidents in 2019 - State-Wise



Map 2.2 Persons Killed in Road Accidents in 2019 - State-Wise



Source: https://morth.nic.in/sites/default/files/RA_Uploading.pdf

MATERIALS AND METHODS

Our project Objective is to check if there is any association between the age and sex of the person died in road fatality. To perform this, we carry out **Pearson's Chi-Squared test.**

On the other hand, to check whether the rate of accident in rainy season differ significantly from that in other seasons. To perform this, we carry out the test of **Equality of Poisson parameters** Since the number of accidents usually follow a Poisson distribution. We will also use a nonparametric test i.e., **Wilcoxon Rank sum test** To examine if the average rate of change of accidents caused in sunny seasons is significantly different from that of the accidents during rainy season for the interval 2019-2020.

PEARSON'S CHI SQUARE TEST :-

Pearson's chi square test is test of independence which assesses whether unpaired observations on two variables, expressed in a contingency table, are independent of each other. (e.g., the age and sex of the person died in road fatality is related to each other or not).

If at the end of the result that we have found is that if p-value is less than the level of significance, we reject the null hypothesis.

TESTING FOR STATISTICAL INDEPENDENCE

The Chi-Square Test of Independence determines whether there is an association between categorical variables (i.e., whether the variables are independent or related). It is a nonparametric test.

This test is also known as: **Chi-Square Test of Association**.

This test utilizes a contingency table to analyze the data. A contingency table (also known as a *cross-tabulation*, *crosstab*, or *two-way table*) is an arrangement in which data is classified according to two categorical variables. The categories for one variable appear in the rows, and the categories for the other variable appear in columns. Each variable must have two or more categories. Each cell reflects the total count of cases for a specific pair of categories.

COMMON USES :-

The Chi-Square Test of Independence is commonly used to test the following: Statistical independence or association between two or more categorical variables.

The Chi-Square Test of Independence can only compare categorical variables. It cannot make comparisons between continuous variables or between categorical and continuous variables.

Additionally, the Chi-Square Test of Independence only assesses *associations* between categorical variables, and cannot provide any inferences about causation.

OUR DATA MUST MEET THE FOLLOWING REQUIREMENTS :-

1. Two categorical variables.
2. Two or more categories (groups) for each variable.
3. Independence of observations.
    - There is no relationship between the subjects in each group.
    - The categorical variables are not "paired" in any way (e.g., pre-test/post-test observations).
4. Relatively large sample size.
    - Expected frequencies for each cell are at least 1.
        - Expected frequencies should be at least 5 for the majority (80%) of the cells

HYPOTHESIS :-

The null hypothesis (H_0) and alternative hypothesis (H_1) of the Chi-Square Test of Independence can be expressed as:

H_0 : "[*Variable 1*] is independent of [*Variable 2*]"
H_1 : "[*Variable 1*] is not independent of [*Variable 2*]"

TEST STATISTIC :-

The test statistic for the Chi-Square Test of Independence is denoted X^2 , and is computed as:

$$\chi^2 = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{(o_{ij}-e_{ij})^2}{e_{ij}}$$

where,

o_{ij} is the observed cell count in the i^{th} row and j^{th} column of the table

e_{ij} is the expected cell count in the i^{th} row and j^{th} column of the table, computed as

$$e_{ij} = \text{row } i \text{ total} * \text{col } j \text{ total grand total}$$

The quantity (o_{ij} - e_{ij}) is sometimes referred to as the *residual* of cell (i, j), denoted r_{ij}.

The calculated χ^2 value is then compared to the critical value from the χ^2 distribution table with degrees of freedom df = (*R* - 1)(*C* - 1) and chosen confidence level. If the calculated χ^2value > critical χ^2 value, then we reject the null hypothesis.

TEST OF EQUALITY OF POISSON PARAMETERS

ASSUMPTIONS

1. The data are counts (discrete) that follow the Poisson distribution.

2. The sample is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

Let x_{1i} (i=1,2,3, $\dots, n_1$) be a random sample from a $Poisson\ (\lambda_1)$ distribution and x_{2i}(i=1,2,3, $\dots, n_2$) be a random sample from $Poisson\ (\lambda_2)$ distribution, The two samples themselves being mutually independent. The null hypothesis we consider is $H_0: \lambda_1 = \lambda_2$. A test for the above hypothesis should be based on the sufficient statistics

$$y_1 = \sum_{i=1}^{n_1} x_{1i}$$

$$y_1 = \sum_{i=1}^{n_2} x_{2i}$$

We consider only those samples of sizes n_1 and n_2 for which $y = y_1 + y_2$ is a constant,(the same as the observed sum of y_1 and y_2. Under H_0, if we denote the common value of the two parameters by λ, then the p.m.f.'s of y_1, y_2 and $y = y_1 + y_2$ are

$$f(y_1) = \frac{e^{(-n_1\ \lambda)}\ (n_1\lambda)^{y_1}}{y_{1!}} \qquad f(y_2) = \frac{e^{(-n_2\ \lambda)}\ (n_2\lambda)^{y_2}}{y_{2!}}$$

$$f(y) = \frac{e^{(-(n_1\ +n\_2)\lambda)}\ ((n_1+n_2)\lambda)^y}{y!}$$

The conditional p.m.f of y_1 for given y is, therefore,

$$f(y_1|y) = \frac{\frac{\exp[-(n_1+n_2)\lambda](n_{1\lambda})^{y_1}(n_{2\lambda})^{y-y_1}}{\{y_1!(y-y_1)!\}}}{\frac{\exp[-(n_1+n_2)\lambda][(n_1+n_2)\lambda]^y}{y!}}$$

$$f(y_1|y) = \binom{y}{y_1}\left(\frac{n_1}{n_1+n_2}\right)^{(y_1)}\left(\frac{n_2}{n_1+n_2}\right)^{y-y_1},$$

i.e. is binomial with parameters y and $\frac{n_1}{(n_1+n_2)}$

This again is free from the unknown nuisance parameter λ.

Denoting the observed values of y_1 and y by y_{10} and y_0, respectively, We consider the conditional p.m.f. $f(y_1|y_0)$ for testing H_0

(a) If the alternative is $H_1: \lambda_1 > \lambda_2$, we compute

$$P[y_1 \geq y_{10} \,|\, y = y_0] = \sum_{y_1 \geq y_{10}} \binom{y_0}{y_1} \left(\frac{n_1}{n_1+n_2}\right)^{(y_1)} \left(\frac{n_2}{n_1+n_2}\right)^{y_0-y_1}$$

And compare it with α for acceptance or rejection of H_0.

(b) If, instead, we are interested in the alternative H: $\lambda_1 < \lambda_2$, we compute

$$P[y_1 \leq y_{10} \,|\, y = y_0] = \sum_{y_1 \leq y_{10}} \binom{y_0}{y_1} \left(\frac{n_1}{n_1+n_2}\right)^{(y_1)} \left(\frac{n_2}{n_1+n_2}\right)^{y_0-y_1}$$

And compare it with α for acceptance or rejection of H_0.

APPROXIMATE TESTS AND CONFIDENCE LIMITS FOR POISSON PARAMETERS

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample from a Poisson distribution with unknown parameter λ. An approximate test or confidence interval for λ can be obtained from the fact that sufficient statistic

$y = \sum_i x_i$ is approximately normally distributed with mean and variance both equal to nλ, provided nλ is sufficiently large.

A test for

$$H_0: \lambda = \lambda_0$$

Will then be given by the statistic

$\frac{(y-n\lambda_0)}{\sqrt{n\lambda_0}}$, Which is approximately normally distributed under H_0

Similarly, The fact that approximately $P\left[-\frac{\tau_\alpha}{2} \leq \frac{y-n\lambda}{\sqrt{n\lambda}} \leq \frac{\tau_\alpha}{2}\right] = 1-\alpha$

Will provide us with confidence limits to λ, the associated confidence coefficient being approximately 1-α.

Again, we may be interested in a comparison among k distributions with unknown parameters $\lambda_i. (i = 1,2, \dots k)$. If y_i $(i = 1,2, \dots, k)$ be the totals, for the k distributions, of independent random samples taken from them, then

$$\sum_i \frac{(y_i - n_i\,\lambda_i)^2}{n_i\lambda_i}$$

Is approximately a χ^2 with df=k. Hence to test for the hypothesis

$$H_0: \lambda_1 = \lambda_1 = \cdots = \lambda_k$$

We could use the statistic

$$\sum_i \frac{(y_i - n_i\,\lambda_i)^2}{n_i\lambda_i}$$

(Approximately a χ^2 with df = k under H_0)

If the common value λ were given. But λ will in most cases be unspecified and will have to be estimated from the data. The maximum-likelihood estimate is $\frac{\sum_{i=1}^{k} y_i}{\sum_{i=1}^{k} n_i}$ If we replace λ by $\frac{\sum_{i=1}^{k} y_i}{\sum_{i=1}^{k} n_i}$, We may still use

$$\sum_i \frac{(y_i - n_i\,\lambda_i\hat{})^2}{n_i\lambda_i\hat{}}$$

For an appropriate test H_0. This is distributed approximately as a χ^2 with df= k-1, the loss of one degree of freedom resulting from estimation of λ by its estimated value.

MANN WHITNEY U TEST (WILCOXON RANK SUM TEST)

The modules on hypothesis testing presented techniques for testing the equality of means in two independent samples. An underlying assumption for appropriate use of the tests described was that the continuous outcome was approximately normally distributed or that the samples were sufficiently large (usually $n_1 \geq 30$ and $n_2 \geq 30$) to justify their use based on the Central Limit Theorem. When comparing two independent samples when the outcome is not normally distributed and the samples are small, a nonparametric test is appropriate.

A popular nonparametric test to compare outcomes between two independent groups is the Mann Whitney U test. The Mann Whitney U test, sometimes called **the Mann Whitney Wilcoxon Test** or **the Wilcoxon Rank Sum Test**, is used to test whether two samples are likely to derive from the same population (i.e., that the two populations have the same shape). Some investigators interpret this test as comparing the medians between the two populations. Recall that the parametric test compares the means (H_0: $\mu_1=\mu_2$) between independent groups.

In contrast, the null and two-sided research hypotheses for the *nonparametric test* are stated as follows:

H_0: The two populations are equal versus

H_1: The two populations are not equal.

This test is often performed as a two-sided test and, thus, the research hypothesis indicates that the populations are not equal as opposed to specifying directionality. A one-sided research hypothesis is used if interest lies in detecting a positive or negative shift in one population as compared to the other. The procedure for the test involves pooling the observations from the two samples into one combined sample, keeping track of which sample each observation comes from, and then ranking lowest to highest from 1 to n_1+n_2, respectively.

USES

Use it when the data do not meet the requirements for a parametric test (i.e. if the data are not normally distributed; if the variances for the two conditions are markedly different; or if the data are measurements on an ordinal scale or for any continuous distribution without the assumption of population distribution). Mann Whitney U test is the nonparametric version of ”two-sample” t-test.

In Mann-Whitney U test we see that how the two sample ranks differ. We count the number of times an xi from sample 1 is greater than a yj from sample 2. This number is denoted by Ux. Similarly, the number of times an xi from sample 1 is smaller than a yj from sample 2 is denoted by Uy. Under the null hypothesis we would expect Ux and Uy to be approximately equal (as under H0 the distributions are same), so $P(xi > yj) = P(xi < yj) = 1/2$. Note that Ux + Uy total pairs (xi , yj) to be compare= n1,n2.

PROCEDURE

1. Mix the two samples and arrange all the observations in order of magnitude.

2. Under each observation, write down X or Y (or some other relevant symbol) to indicate which sample they are from.

3. Under each x write down the number of ys which are to the left of it (i.e. smaller than it); this indicates xi > yj .

4. Under each y write down the number of x s which are to the left of it (i.e. smaller than it); this indicates yj > xi

5. Add up the total number of times xi > yj which is denoted by Ux. Add up the total number of times yj > xi which is denoted by Uy.

6. Count .5 for any two observations which are equal, i.e. xi = yj .

7. the rejection rule:

• If H_1: F_X (1st population, say) $>$ F_Y (2nd population), reject H_0 if U_X < Uα (left tail critical region).α = level of significance. $U_X = no\ of\ times\ xi\ >\ yj$. Uα is the tabular value.

.

• If H_1: F_X (1st population, say) $<$ F_Y (2nd population), reject H0 if $U_Y < U\alpha$ (left tail critical region). Remember

U_Y = no of times yj > xi

.Uα is the tabular value. Same table of Uα.

• If H_1: F_X (1st population, say) = F_Y (2nd population). Find U = min(UX, UY). Table is found for two tailed test (Different table). Reject H_0 if U < Uα.

Wilcoxon Mann Whitney Test An extension of Mann-Whitney test is Wilcoxon Mann Whitney test which is used to test same as Mann Whitney test does. Only difference is in calculation of test statistic.

In Wilcoxon Mann Whitney test one has to find the rank of X observations and Y observations in the combined series and UX and UY are formulated in terms of ranks as follows.

$$U_X = R_1 - \frac{n_1(n_1+1)}{2} \text{ and } U_Y = R_2 - \frac{n_2(n_2+1)}{2}2$$

where R1 = total rank of first sample observations (x) in combined sample and R2 =total rank of second sample observations (y) in combined sample. The rejection rules remains same as in Mann Whitney test.

ANALYSIS OF THE DATA

ANALYSIS USING PEARSON'S CHI SQUARE TEST :-

To check if there is any association between the age and sex of the person died in road fatality for the year 2019

• **Testing of Hypothesis**:

Null Hypothesis:

H0: Sex of the person died is independent of Age – group of the person died.

H1: Sex of the person died depends upon Age – group of the person died.

In the following we arrange our data in the form of a contingency table. The results of the peorson's chi square test of independence are as follows:

CONTINGENCY TABLE:

ACTUAL OBSERVATIONS (year 2019)		AGE GROUP						
		< 18	18-25	25 - 35	35 - 45	45 - 60	60 <	TOTAL
SEX	MALE	2699	14794	17042	12931	7750	2092	57308
	FEMALE	250	632	650	403	218	92	2245
TOTAL		2949	15426	17692	13334	7968	2184	59553

EXPECTED OBSERVATIONS (year 2019)		AGE GROUP						
		< 18	18-25	25 - 35	35 - 45	45 - 60	60 <	TOTAL
SEX	MALE	2837.83	14844.48	17025.05	12831.34	7667.63	2101.67	57308
	FEMALE	111.17	581.52	666.95	502.66	300.37	82.33	2245
TOTAL		2949	15426	17692	13334	7968	2184	59553

- **<u>Test Statistics:</u>**

$X^2 = \sum \{(\text{Observed} - \text{Expected})^2\}/\text{Expected}$

Here,

$(2699 - 2837.83)^2/2837.83 + (14794 - 14844.48)^2/14844.48 + (17042 - 17025.05)^2/17025.05 + (12931 - 12831.34)^2/12831.34 + (7750 - 7667.63)^2/7667.63 + (2092 - 2101.67)^2/2101.67 + (250 - 111.17)^2/111.17 + (632 - 581.52)^2/581.52 + (650 - 666.95)^2/666.95 + (403 - 502.66)^2/502.66 + (218 - 300.37)^2/300.37 + (92 - 82.33)^2/82.33$

=230.3516(very large) > **5.99 (=** χ^2 **0.05;(6-1)(2-1)=** χ^2 **0.05;5 =11.070**

DECISION: So, the null hypothesis is rejected. Therefore, we can say that the sex of persons died in accidents is dependent upon the age-group.

```
R-OUTPUT
Pearson's Chi-squared test

data:  df
X-squared = 230.35, df = 5, p-value < 2.2e-16
```

ANALYSIS USING POISSON TEST

To check whether the rate of accidents in rainy season differ significantly from that in rainy seasons, for the year 2019

WEATHER / ZONE	SUNNY	RAINY
EASTERN	17941	5169
WESTERN	49765	2885
NORTHERN	38737	3632
SOUTHERN	155602	11644
CENTRAL	62504	14247
NORTH-EASTERN	5746	2248

Let x_{1i} (i=1,2,3, ... , $n_{1=}$36) be a random sample from a $Poisson\ (\lambda_1)$ distribution and x_{2i}(i=1,2,3, ... , $n_2 = 36$) be a random sample from $Poisson\ (\lambda_2)$ distribution

The two samples themselves being mutually independent . The null hypothesis we consider is H_0: $rate\ of\ accidents\ in\ sunny\ season_1 = rate\ of\ accidents\ in\ rainy\ season_2$. Against rate of accidents in sunny season> rate of accidents in rainy season for 2019

A test for the above hypothesis should be based on the sufficient statistics

$$y_1 = \sum_{i=1}^{36} x_{1i}$$

$\sum_{i=1}^{36} x_{1i}$=4217+2363+6410+4951+32+52+3131+15014+31536+299+2636+4619+2501+3838+3416+21428+182+17340+35753+35197+1+656+48302+9033+34681+17703+18171+1087+62+4322+394+181+44+130+79+534

$\sum_{i=1}^{36} x_{1i}$=$\sum_{i=1}^{6} No. of\ accidents\ in\ ith\ zone\ in\ sunny\ days$ =17941+49765+38737+155602+62504+5746=330295

$y_1 = \sum_{i=1}^{36} x_{2i}$

$\sum_{i=1}^{36} x_{2i}$=2203+828+1810+328+17+9 +309+1364+1186+4+800+1555+183+148+587+355+48+874+3181+4571+0+143+66+2761+1535+5549+7001+1862+49+1862+148+20+9+91+ 21+48=$\sum_{i=1}^{6} No. of\ accidents\ in\ ith\ zone\ in\ rainy\ days$

=5169+2885+3632+11644+14247+2248=39825.

Rate ratio = $\frac{\sum_{i=1}^{6} No.of\ accidents\ in\ ith\ zone\ in\ sunny\ days\ /n\_1}{\sum_{i=1}^{6} No.of\ accidents\ in\ ith\ zone\ in\ rainyy\ days\ /n\_2} = \frac{330295}{39825.}$=8.293659761

Sample	N	Total Occurrences	Sample Rate
sunny	36	330295	9174.86
rainy	36	39825	1106.25

<u>R-OUTPUT</u>

```
    Comparison of Poisson rates

data:  c(sum(x1), sum(x2)) time base: c(length(x1), 
length(x2))

count1 = 330295, expected count1 = 185060, p-value < 
2.2e-16

alternative hypothesis: true rate ratio is not equal 
to 1

95 percent confidence interval:

 8.207852 8.380541

sample estimates:

rate ratio 

   8.29366
```

```
Comparison of Poisson rates

data:  c(sum(x1), sum(x2)) time base: c(length(x1), 
length(x2))

count1 = 330295, expected count1 = 185060, p-value < 
2.2e-16

alternative hypothesis: true rate ratio is greater 
than 1

95 per

cent confidence interval:

 8.221548      Inf

sample estimates:

rate ratio 8.29366
```

DECISION: Our p-value is less than 0.05 i.e., our H_0 is rejected. **the rate of accidents in rainy season differ significantly from that in rainy seasons, for the year 2019 and that also the rate of accidents in sunny season is more than that of rainy season.**

CAUSES OF MORE ACCIDENTS DURING SUMMER

There appears to be a variety of reasons for more crashes in the summer months. The following list contains common causes of summer road accidents.

TRAFFIC CONGESTION

In summer, there are more travellers on the roadways, and traffic congestion leads to more crashes. The nice summer weather also brings out motorcyclists and bicyclists, and it can be harder for other drivers to see them.

ROAD CONSTRUCTION

Another issue regarding summer driving is the fact that planned transportation construction projects are mostly done during the milder weather, which is typically spring and summer. There are also all those winter weather potholes that need to be repaired.

IMPAIRED DRIVING

Summer is also a time when many people take vacations and celebrate at barbeques where alcohol is often served. Drinking and driving increases during summer months and contributes to an increase in the rate of crashes.

The NHTSA has found that the number of alcohol-impaired drivers increases significantly. Summer crashes involving alcohol impaired drivers is almost double the number of fatal car crashes that occur in all of the other months combined.

DROWSY DRIVING

Being outside in hot weather can also be a factor in the increased number of crashes in summer. The heat can cause dehydration and tiredness. This can reduce attention span, negatively impact judgement, and cause a slower reaction time. All of these can contribute to the increased risk of being involved in a crash.

TOURISTS

Tourism is high during the summer months, especially near vacation destinations, such as beaches, lakes, and campsites. Many tourists are visiting a place for the first time. Their lack of familiarity with the area can cause them to drive while distracted if looking at a navigation device or printout of directions. It can also tend to make them more tentative and less able to negotiate unfamiliar roads, sudden turns, or turnoffs with poor or missing signage. Each of these can make a motorist more prone to driving mistakes.

ANALYSIS USING WILCOXON RANK SUM TEST

To check whether the average rate of change in road accident cases in sunny season for the interval 2019-2020 is significantly different from that of the average rate of change in road accident cases in rainy season.

We have two samples ,sample1 say the number of accidents in sunny season and sample2 be the number accidents during rainy season..

Suppose number of accidents in both the seasons follow two continuous independent populations with distribution functions $F_{sample1}(x)$ and $F_{sample2}(x)$ respectively.

$$n_1 = 36, n_2 = 36$$

sample is 10 or more, we can assume that our sampling distribution is approximately normal. This means wecan use a Z-ratio to calculate the value of p.

We want to test:

$$\boldsymbol{H_0: F_{sample1}(x) = F_{sample2}(x)}$$

$$\boldsymbol{H_0: F_{sample1}(x) > F_{sample2}(x)}$$

H_0:The average rate of change in road accident cases during both season are same for the interval 2019-2020.

H_1 :The average rate of change in road accident cases during sunny season in >The average rate of change in road accidents during rainy season in the interval 2019-2020.

For this situation . We have the rejection rule 1:

• If H_1: F_X (1st population) $>$ F_Y (2nd population), reject H_0 if U_X < Uα (left tail critical region).α = level of significance. $U_X = no\ of\ times\ xi > yj$. Uα is the tabular value.

.We have our test statistic $U_{sample1}$ Which is the count of times where accidents in sunny season > accidents in rainy season for every possible pair.

First, We calculate for the average rate of change in accidents for this We calculate the value of

$\frac{No.of\ accidents\ in\ 2020\ sunny\ season - No.of\ accidents\ in\ 2019\ sunny\ season}{No.of\ accidents\ in\ 2019\ sunny\ season}$ and same for the rainy season$\frac{No.of\ accidents\ in\ 2020\ sunny\ season - No.of\ accidents\ in\ 2019\ sunny\ season}{No.of\ accidents\ in\ 2019\ sunny\ season}$

for each states/UTs

Now, We will form a summarized table in order to calculate the test statistic.

	RATE OF ACCIDENTS IN SUNNY SEASON(2019-2020)	RATE OF ACCIDENTS IN RAINY SEASON(2019-2020)	SAMPLE 1(ACCIDENTS IN SUNNY)	RANK ASSIGNED TO SAMPLE 1	SAMPLE 2(ACCIDENTS IN RAINY)	RANK ASSIGNED TO SAMPLE 2
Andaman & Nicobar Islands	-0.32967033	-0.604166667	-1	1.5	-1	1.5
Andhra Pradesh	-0.144405998	0.574370709	-0.640883978	5	-0.75	3
Arunachal Pradesh	-0.419354839	-0.469387755	-0.491638796	8	-0.66875	4
Assam	-0.354002776	-0.240601504	-0.419354839	13	-0.604166667	6
Bihar	0.051932654	-0.117113028	-0.354002776	19	-0.506756757	7
Chandigarh	-0.491638796	-0.75	-0.32967033	20	-0.469387755	9
Chhattisgarh	-0.067198052	-0.103583062	-0.319686337	22	-0.464480874	10
Dadra & Nagar Haveli	0.3125	0.411764706	-0.302139891	24	-0.45	11
Daman & Diu	-1	-1	-0.294388224	25	-0.430985915	12
Delhi	-0.09522003	-0.66875	-0.290983607	26	-0.409780776	14
Goa	-0.302139891	-0.385113269	-0.279187817	27	-0.395833333	15
Gujarat	-0.244172106	-0.023460411	-0.273569452	28	-0.387661343	16
Haryana	-0.183589522	-0.163987138	-0.264044944	29	-0.385113269	17
Himachal Pradesh	-0.18912435	-0.464480874	-0.244172106	30	-0.384615385	18
Jammu & Kashmir	-0.103699844	0.087837838	-0.240709031	31	-0.325966851	21
Jharkhand	-0.22640711	-0.109903382	-0.235518198	33	-0.303030303	23
Karnataka	-0.190893072	0.234203081	-0.22640711	34	-0.240601504	32
Kerala	-0.319686337	-0.387661343	-0.226292701	35	-0.195400657	38
Lakshadweep						
Madhya Pradesh	-0.178310891	-0.040728059	-0.208841463	36	-0.163987138	44
Maharashtra	-0.240709031	-0.409780776	-0.204545455	37	-0.117113028	47
Manipur	-0.279187817	-0.506756757	-0.190893072	39	-0.109903382	48
Telangana	-0.640883978	-0.45	-0.18912435	40	-0.104938272	49
Meghalaya	-0.204545455	0.333333333	-0.183589522	41	-0.103583062	51
Mizoram	0.561538462	0.417582418	-0.178310891	42	-0.100609756	52
Nagaland	0.009516381	-0.325966851	-0.164556962	43	-0.044293015	56
Odisha	-0.208841463	-0.384615385	-0.144405998	45	-0.040728059	57
Puducherry	-0.290983607	-0.044293015	-0.134114798	46	-0.023460411	58
Punjab	-0.226292701	-0.430985915	-0.103699844	50	0.087837838	61
Rajasthan	-0.164556962	0.523809524	-0.09522003	53	0.234203081	62
Sikkim	-0.235518198	0.692502716	-0.075136336	54	0.333333333	64
Tamil Nadu	-0.134114798	-0.303030303	-0.067198052	55	0.411764706	65
Tripura	-0.264044944	-0.395833333	0.009516381	59	0.417582418	66
Uttar Pradesh	-0.273569452	-0.195400657	0.051932654	60	0.523809524	67
Uttarakhand	-0.294388224	-0.104938272	0.3125	63	0.574370709	69
West Bengal	-0.075136336	-0.100609756	0.561538462	68	0.692502716	70

RESULT DETAILS

H_0:Average Rate of accidents during both season are same in 2019
H_1 :Average Rate of accidents during sunny season in 2019>Average Rate of accidents during rainy season in 2019.

Sample 1
Sum of ranks: 1241.5
Mean of ranks: 35.47
Expected sum of ranks: 1242.5
Expected mean of ranks: 35.5
U-value: 613.5
Expected U-value: 612.5

Sample 2
Sum of ranks: 1243.5
Mean of ranks: 35.53
Expected sum of ranks: 1242.5
Expected mean of ranks: 35.5
U-value: 611.5
Expected U-value: 612.5

Sample 1 & 2 Combined
Sum of ranks: 2485
Mean of ranks: 35.5
Standard Deviation: 85.134

Result 1 - U-value

The U-value is 611.5.

Result 2 - Z-ratio

The Z-Score is -0.00587. The p-value is .49601. The result is not significant at $p < .05$.

alternative hypothesis: true location shift is greater than 0

We reject H_0 .The RATE of accidents that occur in sunny season is is approximately the same as that of RATE of accidents during rainy season.

R-OUTPUT

```
	Wilcoxon rank sum test with continuity correction

data:  x and y
W = 611.5, p-value = 0.507
alternative hypothesis: true location shift is greater than 0
Warning message:
In wilcox.test.default(x, y, "greater") :
  cannot compute exact p-value with ties
```

Result

The average rate of change of road accident cases that occur in sunny season is approximately the same as that of accidents during rainy season for the interval 2019-2020.

CONCLUSION

From the above discussion of tests and results it is concluded that :

the sex of persons died in accidents is dependent upon the age-group.

the rate of accidents during rainy season differ significantly from that during rainy seasons, for the year 2019 and that also the rate of accidents in sunny season is more than that of rainy season

The average rate of change of road accident cases that occur in sunny season is approximately the same as that of accidents during rainy season for the interval 2019-2020.(Since the number of accidents increase/decrease in the same rate for both sunny and rainy season)

We can say that although there is bad road condition in rainy season but the rate of accidents in sunny season is more and there is no rate of change of road accident cases in sunny season and in rainy season. This is mainly due to

Traffic Congestion

Road construction

Impaired driving

Drowsy driving

Tourism

We need to implement strict rule against traffic congestion during summer more.

It takes a heavy financial toll on the economy over and above the mortality and morbidity associated with Road traffic accidents.

Road traffic accidents are predictable and preventable. For it understanding the different factors leading to RTA is a must. Here, In Our analysis We saw that the age-Group and the gender play important role and the weather conditions influence the road fatality.

We need to educate adults (the people of Age-group 18-45) more about the adverse effects of road accidents deaths and Injury and that also to the gender “male” . About three quarters of all road traffic deaths occur among young males under the age of 25 years who are almost 3 times as likely to be killed in a road traffic crash as young females.
Strict implementation of traffic rules and stringent punishments alone will not solve the persisting problem. Change in the mind set of riders and drivers and road users realizing their responsibilities alone will bring about a change among adults.

REFERENCE

https://www.jcdelaw.com/2021/05/20/car-accidents-summer/

Fundamentals of Statistics 1 by A.M GUN, M.K.GUPTA, B.DASGUPTA (Book)

NOTES ON NON PARAMETRIC STATISTICAL TESTS-1 BY Dr. SARAN ISHIKA MAITI

https://morth.nic.in/sites/default/files/RA_Uploading.pdf By Government of INDIA MINISTRY OF ROAD TRANSPORT AND HIGHWAYS

http://www.bro.gov.in/WriteReadData/linkimages/5768690382-14.pdf

APPENDIX

```
> x=c(17941,49765, 38737, 218106,5746)
> y=c(5169,2885,3632, 25891,2248)

> poisson.test(c(sum(x),sum(y)),c(length(x),length(y)), alternative = "greater")
#whether the rate of accidents increases during sunny season than the rainy season
> poisson.test(c(sum(x),sum(y)),c(length(x),length(y)), alternative = "greater")#accidents in rainy season is more than accidents in sunny season

>x1=c(4217,2363,6410,4951,32,52,3131,15014,31536,299,2636,4619,2501,3838,3416,21428,182,17340,35753,35197,1,656,48302,9033,34681,17703,18171,1087,62,4322,394,181,44,130,79,534)
>x2=c(2203,828,1810,328,17,9,309,1364,1186,4,800,1555,183,148,587,355,48,874,3181,4571,0,143,2761,1535,5549,7001,66,162,49,1862,148,20,9,91,21,48)
>poisson.test(c(sum(x1),sum(x2)),c(length(x1),length(x2)), alternative = "greater")#accidents in sunny season is more than accidents in rainy season
```

```
>df=read.csv(file.choose(),header = T) #accident_age_sex_ DATA

>df
>View(df)
>df=df[,-c(1)]
>View(df)
>chisq<-chisq.test(df)
>chisq
```

```
x=c(-0.32967033,-0.144405998,-0.419354839,-0.354002776,0.051932654, -0.491638796, -0.067198052, 0.3125,-1,-0.09522003, -0.302139891,-0.244172106,-0.183589522, -0.18912435,-0.103699844, -0.22640711,-0.190893072,-0.319686337,-0.178310891,-0.240709031,-0.279187817,-0.640883978,-0.204545455,0.561538462,0.009516381,  -0.208841463, -0.290983607,-0.226292701,-0.164556962,-0.235518198,-0.134114798,-0.264044944, -0.273569452,-0.294388224,-0.075136336)#rate 2019-2020 sunny
y=c(-0.604166667, 0.574370709, -0.469387755, -0.240601504,-0.117113028,-0.75, -0.103583062,0.411764706, -1,-0.66875,-0.385113269,-0.023460411, -0.163987138, -0.464480874,0.087837838,-0.109903382,0.234203081,-0.387661343,-0.040728059,-0.409780776, -0.506756757, -0.45, 0.333333333, 0.417582418, -0.325966851, -0.384615385,-0.044293015,-0.430985915, 0.523809524,  0.692502716, -0.303030303, -0.395833333,-0.195400657,-0.104938272, -0.100609756)#rate 2019-2020 rainy
wilcox.test(x,y,"greater")
```