iJDASE International Journal of Data Analytics and Smart Education

iJDASE | ISSN: 2736-979X | Vol. 1 No. 1 (2025)
https://doi.org/10.64817/3c6nqc10

PAPER

Analytics on the Causal Factors Associated with Age, Injuries, and Fatalities in the USA (2001 – 2020)

Toluwalase Janet Akingbade[1,6], **Oluwafemi Samson Balogun**[2,6], **Donald Douglas Atsa'am**[3,6], **Samuel Nii Odoi Devine**[4,6], **Temidayo Oluwatosin Omotehinwa**[5,6]

[1]Department of Industrial Mathematics, Covenant University, Ota, Ogun State, Nigeria

[2]School of Computing, University of Eastern, FI-70211. Finland

[3]Department of Computer Science, Joseph Sarwuan Tarka University Makurdi, Nigeria

[4]Department of Information and Communication Technology & Mathematics, Presbyterian University, Abetifi, Ghana

ABSTRACT

Injuries and fatalities continue to pose major public health challenges across all age groups. This study explores how age, cause of incident, and time influence injury and death rates in the United States from 2001 to 2020, using data from the CDC. By constructing injury and death rates per 100,000 population and applying Poisson and Negative Binomial regression alongside Chi-square analysis, we uncover significant trends and associations. Findings show that age and cause are strong predictors of injury rates—with falls and motor vehicle-related causes emerging as major contributors. Age group 6, in particular, is linked with notably higher injury rates. While injury rates initially appeared to increase over time in Poisson models, this trend was less pronounced when overdispersion was properly addressed using a Negative Binomial model. Conversely, death rates declined steadily over the two decades, especially in earlier years, suggesting improvements in healthcare and prevention.

However, this study is not without limitations. It relies on aggregated national data, which may overlook local and contextual factors such as socioeconomic status or healthcare access. Future studies should consider more granular data, individual-level trajectories, and a broader range of explanatory variables. Investigating the post-2020 period, especially in the context of COVID-19, could also offer valuable insights.

Overall, this research emphasizes the need for targeted, age- and cause-specific strategies to reduce injury and death, helping to shape policies that protect vulnerable populations and enhance public health outcomes.

Toluwalase Janet Akingbade, Oluwafemi Samson Balogun, Donald Douglas Atsa'am, Samuel Nii Odoi Devine, Temidayo Oluwatosin Omotehinwa. (2025). Analytics on the Causal Factors Associated with Age, Injuries, and Fatalities in the USA (2001 – 2020). *International Journal of Data Analytics and Smart Education (iJDASE),* 1(1) pp. 55-83 https://doi.org/10.64817/3c6nqc10

Article submitted: 27.02.2025 Revision uploaded: 16.06.2025 Final Acceptance: 11.07.2025

[5]Department of Mathematics and Computer Science, Federal University of Health Sciences, P.M.B. 145, Otukpo, Nigeria

[6]Centre for Multidisciplinary Research and Innovation (CEMRI), Finland.

toluwalase.akingbade@covenantuniversity.edu.ng, samsb@uef.fi, donatsaam@alumni.emu.edu.tr, samuel.nodevine@presbyuniversity.edu.gh, temidayo.omotehinwa@fuhso.edu.ng

KEYWORDS
Causal factors, injuries, fatalities, USA, analytics, rates

1 INTRODUCTION

There are different factors responsible for human injuries and deaths. These factors vary by geographical location and season. Residents of earthquake- or flood-prone areas are at risk of injuries and deaths related to these natural disasters. Similarly, those in mountainous areas are at risk of falls resulting from hiking and other outdoor activities. Age is a very important factor to consider when examining the causes of human injuries and death. Different age groups are susceptible to different categories of injuries and fatalities. For example, older adults are more susceptible to injuries from falls [1, 2], while younger adults are more at risk of poisoning injuries resulting from the use of illicit drugs, alcohol, and so on [3].

Poisoning is an important factor when analyzing human injuries and fatalities statistics. It can be categorized into two: intentional and non-intentional. The accidental ingestion of harmful substances such as chemicals and medications which can easily occur among children is an example of unintentional poisoning. Among young adults and middle-aged individuals, a deliberate overdose of drugs for some momentary benefit or suicide are common intentional poisoning. In 2016, an estimated $812.5 million was the total cost of poisoning in British Columbia [3]. Of this total, 84% represents the cost of unintentional injuries suffered by young adults and middle-aged individuals between the age of 25-64 [3]. In a study carried out in North Carolina by Shiue et al., [4], it was revealed that 895 of the drug overdose deaths recorded between 2015 and 2019 were unintentional and only 2% of the drug overdose deaths were intentional.

Incidents such as accidents, occupational hazards, and home fires involving flammable substances are some of the contributing factors to fire and burn injuries. These could lead to either disabilities or fatalities. Age is a key contributing factor to fire and burn incidents. Children are more likely to spill scalding liquids on their bodies, and older adults with cognitive impairments and mobility challenges are at a higher risk of fire accidents. According to the World Health Organization Global Burn Registry Report [5], the highest cases of burns occurred among children aged 0 to 4, with flame and hot liquid steam or gas being the leading cause. Of the 9,278 cases reported in the global registry, 1,751 fatalities were recorded [5].

No age group is at risk of transport accidents such as air crashes, motor vehicle accidents, or boat and cycling mishaps. About 1.3 million deaths are recorded every year as a result of road traffic accidents, and more than 50% of these deaths are among vulnerable road users such as pedestrians and cyclists [6]. Some of the factors responsible for road traffic accidents include season and weather conditions [7, 8], over-speeding, and driving under the alcoholic influence [9, 10]. According to Chen et al. [11], it is estimated that between 2015 and 2030, the cost of road traffic accidents on the world economy would be about US$1·8 trillion.

The use of sharp objects, machinery, and other tools can lead to cut and pierce injuries and fatalities. These injuries can lead to nerve damage, extreme bleeding, and permanent disabilities due to amputations [12]. Occupational hazards in industries such as construction and manufacturing predispose working adults to risks of cut and pierce injuries and fatalities. Children and young adults are also at risk of accidental cuts and pierce injuries, which could result in fatalities. There are also cut and pierce injuries, which are products of Intimate Partner Violence (IPV) [13]. The study by Khurana et al., [13] which analyzed the US Emergency Department injury surveillance data, revealed that cut injuries, lacerations, and injuries to the upper limb and arm regions are more common in males in IPV than females. IPV-related injuries in men are known to be underreported due to factors such as stigma, fear, and societal expectations. Understanding the associated injury pattern will assist healthcare providers in proactively identifying and assessing the cause of abuse in men.

Falls have been established as one of the leading causes of injury, particularly among older adults aged 65 years and above [1]. According to WHO [14], about 684,000 deaths resulting from falls are recorded globally every year. Lytras et al. [15] identified poor lighting conditions, psychotropic drugs, dizziness, and lack of safety standards in homes as some of the factors responsible for falls in the elderly. Another group highly susceptible to injury due to falls is children. Children are ‘risk takers’, an attribute they exhibit as they curiously explore their environment. Inadequate supervision could lead to the risk of injury due to uncontrolled physical activities. The impact of falls in both adults and children ranges from moderate to severe injuries such as head trauma, hip and femur fractures, and bruises, and in extreme cases, the injuries could lead to death. In some other cases, these injuries can inflict permanent disabilities. In 2020, the age-adjusted mortality rate due to falls was 10.843 deaths per every 100, 000 individuals in China [16]. The WHO put the global DALYs (disability-adjusted life years) due to falls at 38 million annually [14]. This implies that, beyond premature death, a substantial number of years are lived with disability. This underscores the long-term impact on the affected individuals, the family, and society.

Comprehensive analytics of the causal factors associated with injuries and fatalities, to age, will give significant insights into the trends and patterns exhibited by these disasters. The insights gained will assist in top-level health management decision-making, preventive strategies, and policy formulations with a significant impact on public health. This study is, therefore, aimed at the explorative analytics of factors such as poisoning, fall, fire/burn, cut/piercing, and transportation injuries to age and their impact on injuries and fatalities in the United States of America (USA).

2 RELATED LITERATURE

Several studies have been conducted to investigate causal factors associated with age, injuries, and fatalities. The causes of injuries and fatalities, preventable and non-preventable, span from motor accidents, falls, poisoning, fires and burns, natural disasters, among others. A review of existing literature shows that abundant studies is using various statistical techniques to investigate the factors associated with human injuries and fatalities. Some of those studies are reviewed in this section.

Li et al. [17] conducted a study in which they analyzed the impact of illegal driver behaviors on road traffic accidents in China. They used correlation and partial least square regression statistical analyses to identify the correlation between different factors and road accidents. They found that behaviors such as drunk driving, speeding, and illegal parking were significant risk factors associated with accidents. These results emphasize the importance of addressing these behaviors to reduce accidents and fatalities on the road. Although the study focused on China, its findings can offer valuable insights into a similar situation in other parts of the world.

In a study by Lerdsuwansri et al. [18], road traffic injuries in Thailand were investigated using the Conway-Maxwell-Poisson regression model. The researchers analyzed factors such as road type, road surface, road section, and festive months to determine their association with injury counts per accident. The findings reveal that road type, road section, and festive months significantly impact the number of injuries sustained in accidents. These results emphasize the need for targeted interventions to improve road safety during specific periods and on specific road types.

Rathnayake et al. [19] employed correlation analysis and kernel density estimation in a study done in Sri Lanka to evaluate the association between land use and land cover change (LULCC) and instances of human-elephant conflict (HEC). Their findings demonstrated a link between HEC incidents and land use change events. It also discovered distinct spatial patterns of human deaths, injuries, elephant deaths, and property damage concerning HEC hotspots. These findings emphasize the importance of long-term land management policies and tailored interventions to reduce human-wildlife conflicts. Similarly, Neupane et al. [20] investigated the temporal and spatial patterns of HEC in Nepal. The researchers evaluated historical records using regression and chi-squared testing to uncover patterns and factors leading to HEC events. According to their findings, the intensity of HEC was highest along the migratory path in the eastern Indo-Nepal border region. They also noticed an upsurge in conflict incidents at certain seasons and times of the year. To resolve human-elephant conflicts, the study suggested mitigation methods such as enhanced fencing, corridor development, and resource reallocation.

In Greece, Koutras et al. [21] undertook a statistical investigation of lightning-related deaths, injuries, property damage, and fires. The researchers evaluated the association between lightning strikes and casualties using statistical methods such as correlation analysis and indicator computation. They evaluated data from a specified period to determine the relationship between lightning-related occurrences and other characteristics such as geographic regions and anti-lightning protective measures. The study emphasized the need to take necessary precautions and promote public awareness in order to limit the hazards connected with lightning strikes.

Turner et al. [22] conducted a pilot study in the United States on the causes and rates of mortality among college/university students in America from the self-selected ages of 18 to 24 years. The study was conducted in 2010, during the summer and early fall, using an online survey approach, with 42 questions administered. A total of 157 colleges participated in the survey, with 254 deaths reported. Aspects investigated in the survey included age, gender, various school characteristics, and sources of information about student mortalities. The research equally analyzed the sample and national enrolment and characteristics for the 2009/2010 academic year. Excel 2007 and SPSS v19 were used to manage and analyze the data. Their results indicated that among college students, unintentional accidental deaths were the main causes of mortality, considering both vehicular and non-traffic cases. Further, alcohol-related vehicular and non-traffic injuries were observed to be lower, and suicide was a factor in the death cases recorded.

Halabi et al. [23], in their research, investigated the primary factors of fall-related accidents in the construction industry involving 23057 cases covering 20 years. The secondary data obtained for the study were from the Occupational Safety and Health Administration (OSHA) database from January 2000 to

August 2020. In analysing the data to identify trends in accident cases, frequency analysis was performed. Further, correlation analysis was performed to find the association between the factors and degree of injuries, and logistic regression was used to develop a fatality predictive model for diagnosing accidents and to aid in safety preparation and risk monitoring. One of their findings indicated that most of the fall accidents were among older workers. The researcher realized that their prediction model was able to correctly diagnose the degree of injury outcome by 77.7% based on selected predictors that caused the accident.

In a study conducted in Finland, Malin et al. [24] investigated the differences in the factors associated with pedestrians suffering serious injuries or fatalities (MAIS3+) and the extent of their prevalence. The research examined accident data from 2014 to 2017, comprising all motor vehicle pedestrian accidents and exposure data from the 2016 national travel survey in Finland. The data included a total of 285 killed or seriously injured pedestrians involved in 281 accidents. Factors analyzed included age, gender, and others. Additionally, between pedestrian serious injuries and fatalities, the differences observed were in the place of accident, area and municipal types, vehicle type, speed limit, road conditions, and lighting conditions. The chi-square test was employed to statistically analyse significant differences between factors and binomial regression models were applied to evaluate the relative contributions of relevant variables. It was further observed that for pedestrians who suffered serious injuries and fatalities, the rate of occurrence was five times higher for those aged 75 years and over, with males having approximately 40% higher rates than females.

Wang et al. [25] applied an ordered logistic regression technique to assess the nexus between marine accident severity and corresponding contributory factors. The data used in the study were extracted from marine accident investigation reports obtained from different agencies. The data covers a period of 9 years from 2010–2019 of 1207 accidents involving a total of 1294 ships. Factors considered in the study covered areas related to the accident types, ship conditions, human elements, and environmental conditions such as fire/explosion, sinking, ship age, seafarer’s certificate, sea experience, visibility, and water death, among others. The study revealed that there is a positive correlation between the severity level of marine-related accidents and sinking accidents, heavy seas, good visibility, strong current, strong wind, and/or being far away from the port. Furthermore, the level of severity is higher for ships that do not have complete or valid seafarers’ certificates and/or are more than 30 years old. Pearson correlation and covariance tests were employed for statistical analysis to identify the factors that covariate with other factors.

Taylor et al. [26] in their research studied the vulnerability of the aged to fire-related injury without cases of fatalities. Their study explored a dataset of 279 accidental home-related fire injuries over 10 years in a period in Merseyside, England, from the UK Fire and Rescue Service between April 2011 and April 2022. The study investigated the extent of vulnerability to risk factors associated with fire injuries, considering age groups, gender, alcohol consumption, deprivation, mobility, occupancy level, and while attempting to fight fire. They reported that fire-related injuries were higher among older persons in single occupancy housing and in more deprived settings, with the main causes being cooking-related, smoking-related, or heating-related injuries. Frequency analysis and Spearman's rank-order correlation coefficient for correlation analysis were used to analyze the patterns and trends based on the factors considered in the non-fatal fire injuries.

2.1 Significance of the Work

It gives an insight into the complicated patterns of relations between age, injury, and mortality in the United States over the period from 2001 to 2020, using the comprehensive dataset of the CDC. The use of various methods of analysis in this research work helped to project specificpyз insights into the relationships among demographic variables concerning age, causes of injuries, and fatalities. It is through this understanding that targeted public health policies and interventions would be formulated to reduce the incidence of injuries and fatalities across different age groups.

2.2 Contribution of the Work

The multi-faceted methodological approach for the research involves descriptive statistics, poison and negative binomial regression analysis and Chi-square analysis for specific hypotheses testing. Methods decode complex trends and relationships, demonstrating important associations between age groups and causes of injuries and deaths. The findings indicate that falls are one of the leading causes of both injuries and death. Injuries and death have most of their variability accounted for by age and causes. This research thus provides key evidence to policymakers, public health officials, and safety advocates for every age brackets in which to draft appropriate strategies and interventions to improve safety and reduce risk.

3 METHODOLOGY

This section provides a detailed description of the dataset as well as the methods of analysis used in this study..

3.1 Data Preparation or Variable Construction

To account for population differences across age groups and years, injury and death counts were converted into rates per 100,000 population. This was done by dividing the raw counts by age-group-specific population estimates from the CDC and multiplying by 100,000. Additionally, a time variable (year) was included in all regression models to capture time trends and control for broader shifts over the 2001–2020 period, such as the rise in opioid-related injuries and deaths post-2010.

Injury Rate = (Number of Injuries / Age Group Population) × 100,000

Death Rate = (Number of Deaths / Age Group Population) × 100,000

3.2 Description of the Research Data and Source

The dataset used for this study is secondary data consisting of various causes and the number of deaths and injuries on a yearly basis in the USA (2001 – 2020) are created by the Centre for Disease and Control (CDC) USA. Causes of deaths/injuries include poisoning, falls, fire/burn, and so on. The number of deaths/injuries for each cause is grouped by various age ranges of the victims: <1 year, 1-4 years, and so on.

3.3 Detailed Data Analysis

We used a combination of descriptive statistics, poisson regression analysis, negative binomial regression analysis and Chi-square analysis to analyze the data collected in this study.

Regression analysis is a statistical method that is used to investigate the connection between one or more independent variables and a dependent variable. This can be done in several different ways. This study will make use of regression analysis to investigate the relationship between age and cause of death (both of which are considered independent variables) and the number of deaths and injuries (which is considered the dependent variable). In this study, poisson and negative binomial regression analysis was used.

3.3 Piosson Regression Model

Poisson regression is a type of regression analysis used to model count data, where the outcome variable represents the number of occurrences of an event within a fixed interval of time or space. It is particularly useful when dealing with data that follow a Poisson distribution, which is a probability distribution commonly used to model counts of events over a fixed interval. In Poisson regression, it is assumed that the dependent variable Y, the number of occurrences of an event, has a Poisson distribution given the independent variables $X_1, X_2, \dots, X_p$.

$$P(Y = y) = \frac{e^{-\mu}\mu^{y}}{y!}, y = 0,1,2 \quad (1)$$

where $E(Y) = \mu \ and\ Var(Y) = \mu$. This is called the equi-dispersion property of the Poisson distribution. The log of the mean μ is assumed to be a linear function of the independent variables, that is,

$$In\mu = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad (2)$$

where $Y \sim p(\mu)$ or equivalently,

$$\mu = e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)} \quad (3)$$

This is the model for analyzing count data.

Sometimes, the response may be in the form of events of a certain type that occur over time, space, or some other index of size. In this situation, it is often relevant to model the data as the rate at which events occur. When a response count Y has an index (such as population size) equal to t, the sample rate of occurrence is Y/t. The expected value for rate is μ/t. Thus, for analysis rate data, the model can be written as:

$$In\left(\frac{\mu}{t}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad (4)$$

This model has an equivalent representation as,

$$In\ \mu - In\ t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad (5)$$

The adjustment term $-In\ t$, on the left-hand side of the equation is called an *offset.* Poisson regression is widely used in various fields such as epidemiology, finance, social sciences, and biology, where count data frequently occur. It is often employed when the response variable represents counts of events, such as the number of accidents at an intersection, the number of customer arrivals at a store, the number of defects in a manufacturing process, etc.

3.3.1 Estimation of Parameters Using Maximum Likelihood Estimation (MLE)

In Poisson regression, the model parameters are estimated using the Maximum Likelihood Estimation (MLE) method. This approach essentially asks the question: "What values of the regression coefficients would make the observed data most likely?" In other words, MLE helps us find the set of parameters that best explain the pattern in the data. The following section discusses how this estimation process works specifically in the context of Poisson regression. Let μ_i be the mean of the ith response, for $i = 1,2, \dots, n$. Since the mean response is assumed to be a function of a set of explanatory variables. $X_1, X_2, .., X_n$, the notation $\mu(X_i, \beta)$ is used to denote the function that relates to the mean response μ_i and X_i where the explanatory variables for case i and the regression coefficients (β) are taken into account. Considering the Poisson regression model in the following form:

$$\mu_i = \mu(X_i, \beta) = e^{(X_i, \beta)} \tag{6}$$

From Poisson distribution

$$P(Y, \beta) = \frac{[\mu(X_i, \beta)]^Y e^{-\mu(X_i, \beta)}}{Y!} \tag{7}$$

The likelihood function is denoted by

$$L(Y; \beta) = \prod_{i=1}^{N} P(Y; \beta) = L(Y; \beta) = \prod_{i=1}^{N} \frac{[\mu(X_i, \beta)]^Y e^{-[(X_i, \beta)]}}{Y!}$$

$$L(Y; \beta) = \frac{\left\{\prod_{i=1}^{N} [(\mu(X_i, \beta))]^Y e^{\left[-\sum_{i=1}^{N} \mu(X_i, \beta)\right]}\right\}}{\prod_{i=1}^{N} Y!} \tag{8}$$

The next step is to take the natural logarithm of the likelihood function described above. After that, we differentiate the resulting expression with respect to β, and then set the derivative equal to zero to find the estimates. This gives us the log-likelihood function, which is expressed as follows:

$$InL(Y_i, \beta) = \sum_{i=1}^{N} [Y_i In[\mu(X_i, \beta)] - \mu(X_i, \beta) - In(Y_i!)] \tag{9}$$

$$\frac{\partial}{\partial \beta} [InL(Y; \beta)] = 0$$

3.4 Overdispersion Test

Poisson regression assumes equidispersion, meaning that the mean and variance of the outcome are expected to be roughly equal. However, in practice, this assumption is often not met. To check for this, we look at the deviance divided by degrees of freedom (deviance/df) or the Pearson chi-square divided by degrees of freedom (chi-square/df). If this value is greater than 1, it suggests overdispersion—the variance is larger than expected. If the value is less than 1, it points to underdispersion, where the variance is smaller than the mean.

3.5 Negative Binomial Regression Model

The negative binomial regression model is developed by modifying the Poisson regression model in a way that allows for more flexibility in handling overdispersion. Specifically, it is reformulated so that,

$$In\mu = \beta_0 + \beta_i X_i + \varepsilon_i$$

The negative binomial regression distribution has the form:

$$P(Y = y) = \frac{\Gamma\left({}^{1}/_{\alpha} + y\right)}{\Gamma\left({}^{1}/_{\alpha}\right)y!}\left[\frac{{}^{1}/_{\alpha}}{\left({}^{1}/_{\alpha}\right) + \mu}\right]^{{}^{1}/_{\alpha}}\left[\frac{\mu}{\left({}^{1}/_{\alpha}\right) + \mu}\right]^{y}$$

Where $\Gamma(.)$ is a gamma function. This results in the likelihood function.

$$L(Y_i) = \prod_{i=1}^{N}\frac{\Gamma\left({}^{1}/_{\alpha} + y\right)}{\Gamma\left({}^{1}/_{\alpha}\right)y!}\left[\frac{{}^{1}/_{\alpha}}{\left({}^{1}/_{\alpha}\right) + \mu}\right]^{{}^{1}/_{\alpha}}\left[\frac{\mu}{\left({}^{1}/_{\alpha}\right) + \mu}\right]^{y}$$

Maximum likelihood estimation is the method used to estimate the parameters in a negative binomial regression model. Just like in Poisson regression, the interpretation of the regression coefficients remains the same, they show how the expected count changes with each predictor variable.

3.6 Chi Square Test

The Chi-square test is a statistical test that is used to determine whether or not there is a significant association or relationship between two categorical variables. This test is used to determine whether or not there is a significant association or relationship between two categorical variables. The chi-square test will be utilized in this research to analyze the correlation between age, which is one categorical variable, and the cause of death, which is another categorical variable. The purpose of this test is to assist in determining whether or not there are discernible differences in the pattern of causes of death across a variety of age groups. Calculating an observed chi-square statistic and comparing it to an expected distribution is what the chi-square test does. This comparison is based on the assumption that the variables are independent. If the observed value of chi-square is higher than the critical value, this indicates that there is a significant association between the variables being studied. Data analysis was conducted using the SPSS Statistical Package (Version 28) and STATA 12.

4 RESULTS AND DISCUSSION

This section presents the results of the descriptive statistics, poisson regression analysis, negative binomial regression analysis and Chi-square analysis to analyse this study.

Table 1: Descriptive Statistics

Variables	**Frequency**	**Percentage**
Age group		
<1 years	50	10.2
1-4 years	55	11.2
5-9 years	21	4.3
10-14 years	24	4.9
15-24 years	45	9.2
25-34 years	40	8.2
35-44 years	42	8.6
45-54 years	45	9.2
55-64 years	42	8.6
65-74 years	45	9.2
75-84 years	40	8.2
85+ years	40	8.2
Causes of Deaths/Injuries		
Cut/pierce	1	0.2
Fall	236	48.3
Fire/Burn	43	8.8
Other transport	31	6.3
Poisoning	178	36.4

Table 1 summarizes the descriptive statistics used. The sample consists of 489 participants. Among the causes of deaths/injuries: 236(48.3%) were Fall, 178(36.4%) Poisoning, 43(8.8%) Fire/burn, 31(6.3%) Other transport and 1(0.2%) Cut/Pierce. The Age group consist of 55(11.2%) for 1-4 years, 50(10.2%) less than 1 year, 45(9.2%) for 15-24, 45-54 and 65-74 years, 42(8.6%) for 35-44 and 55-64 years, 40(8.2%) for 25-34, 75-84 and 85+ years, 24(4.9%) for 10-14 years and 21(4.3%) for 5-9 years.

Table 2: Result from Poisson Regression Analysis (using Injury rate, Agr group, causes and year)

```
. poisson injury_rate b2i.Agegrp b3i.Causes b19.year

Poisson regression                                  Number of obs    =
490
                                                    LR chi2(34)      =
51429.40
                                                    Prob > chi2      =
0.0000
Log likelihood = -4464.8335                         Pseudo R2        =
0.8521

------------------------------------------------------------------------------
------
 injury_rate |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------+----------------------------------------------------------------
------
      Agegrp |
          1  |   .2907703   .0214332    13.57   0.000      .248762
.3327785
          3  |   .3889599    .020939    18.58   0.000     .3479201
.4299997
          4  |   .3467222   .0211744    16.37   0.000     .3052211
.3882233
          5  |   .3379228   .0211604    15.97   0.000     .2964491
.3793965
          6  |   .4519667   .0206965    21.84   0.000     .4114024
.4925311
          7  |   .0629597    .023655     2.66   0.008     .0165967
.1093228
          8  |   .3565489    .021091    16.91   0.000     .3152113
.3978865
          9  |    .182842   .0217662     8.40   0.000      .140181
.2255029
         10  |   .2930453   .0213719    13.71   0.000     .2511572
.3349335
         11  |   .1520193   .0219829     6.92   0.000     .1089335
.195105
         12  |  -1.566028   .0375713   -41.68   0.000    -1.639666   -
1.492389
             |
      Causes |
          1  |  -.3910903   .7144838    -0.55   0.584    -1.791453
1.009272
          2  |   2.599717   .0467521    55.61   0.000     2.508084
2.691349
          4  |  -.1089099   .0652196    -1.67   0.095    -.2367381
.0189182
          5  |   .6620951   .0483177    13.70   0.000     .5673941
.7567961
             |
        year |
          0  |   .1334315   .0266473     5.01   0.000     .0812038
.1856592
          1  |   .0773386   .0271383     2.85   0.004     .0241486
.1305286
          2  |    .115286   .0267563     4.31   0.000     .0628447
.1677274
          3  |   .1266422   .0266034     4.76   0.000     .0745005
```

```
.1787839
          4  |   .1096327   .0267237     4.10   0.000     .0572553
.1620101
          5  |   .1073363   .0267266     4.02   0.000      .054953
.1597195
          6  |   .1101467   .0268017     4.11   0.000     .0576164
.162677
          7  |    .169395   .0264215     6.41   0.000     .1176098
.2211801
          8  |   .1727838   .0263273     6.56   0.000     .1211833
.2243843
          9  |   .2160337   .0260667     8.29   0.000      .164944
.2671235
         10  |   .2438632   .0259729     9.39   0.000     .1929572
.2947691
         11  |   .2072112   .0261905     7.91   0.000     .1558789
.2585436
         12  |    .181194   .0267159     6.78   0.000     .1288317
.2335562
         13  |    .236866    .026014     9.11   0.000     .1858796
.2878525
         14  |   .2703881   .0257727    10.49   0.000     .2198744
.3209017
         15  |   .2728665   .0258196    10.57   0.000      .222261
.323472
         16  |   .1990959   .0259825     7.66   0.000     .1481712
.2500206
         17  |   .1677576   .0265648     6.32   0.000     .1156915
.2198238
         18  |   .1344863   .0262923     5.12   0.000     .0829543
.1860183
             |
       _cons |   2.463299   .0531377    46.36   0.000     2.359151
2.567447
------------------------------------------------------------------------
------

. estat gof

         Deviance goodness-of-fit =   6061.292
         Prob > chi2(455)         =     0.0000

         Pearson goodness-of-fit  =   6188.428
         Prob > chi2(455)         =     0.0000

. estat ic

------------------------------------------------------------------------
-----
       Model |     Obs    ll(null)   ll(model)      df          AIC
BIC
-------------+----------------------------------------------------------
-----
           . |     490   -30179.53   -4464.834      35     8999.667
9146.471
------------------------------------------------------------------------
-----
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Source: Personal Collection

To start with, Several age groups have statistically significant positive coefficients ($p < 0.001$), meaning these groups are associated with higher injury rates compared to the reference group. For example: Age group 6 has the highest positive coefficient (0.452), suggesting a strong association with higher injury rates. Age group 12 shows a large negative coefficient (-1.566), indicating significantly lower injury rates than the reference.

Besides, Cause 2 (likely a major external factor like "Motor Vehicle") has a very strong positive effect on injury rates (coefficient = 2.60, $p < 0.001$). Cause 1 is not significant ($p = 0.584$), and Cause 4 is marginal ($p = 0.095$). Cause 5 also has a significant positive effect.

Furthermore, all year variables are statistically significant ($p < 0.01$), with coefficients generally increasing over time. This indicates a positive trend in injury rates from 2001 to 2020, possibly reflecting broader societal changes, such as increased reporting or shifts in public health risk factors.

Pearson $chi^2/df = 6188.43 / 455 \approx 13.6$

Deviance $chi^2/df = 6061.29 / 455 \approx 13.3$

These values are much greater than 1, indicating significant overdispersion in your model. This violates the core Poisson assumption that the mean equals the variance (see table 2)

Table 3: Result from Negative Binomial Regression Analysis (using Injury rate, Agr group, causes and year)

```
  nbreg injury_rate b2i.Agegrp b3i.Causes b19.year

Negative binomial regression                            Number of obs     =
490
                                                        LR chi2(34)       =
1006.46
Dispersion      = mean                                  Prob > chi2       =
0.0000
Log likelihood = -2337.8055                             Pseudo R2         =
0.1771

------------------------------------------------------------------------------
------
 injury_rate |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------+----------------------------------------------------------------
------
      Agegrp |
          1  |   .1035617   .1210794     0.86   0.392    -.1337496
.340873
          3  |   .6154614   .1211272     5.08   0.000     .3780564
.8528664
          4  |    .700691   .1246372     5.62   0.000     .4564066
.9449753
          5  |   .6818577   .1241465     5.49   0.000      .438535
.9251804
          6  |   .7323095   .1221438     6.00   0.000     .4929121
.9717069
```

7	.0078204	.1389109	0.06	0.955	-.26444	.2800808
8	.4545886	.1230554	3.69	0.000	.2134044	.6957727
9	-.0062277	.1203836	-0.05	0.959	-.2421753	.2297198
10	-.0507684	.1220038	-0.42	0.677	-.2898914	.1883546
11	-.2751616	.1220164	-2.26	0.024	-.5143093	-.0360139
12	-1.684142	.1292122	-13.03	0.000	-1.937393	-1.43089
Causes						
1	-.3155533	.8529399	-0.37	0.711	-1.987285	1.356178
2	2.52138	.1044067	24.15	0.000	2.316747	2.726014
4	-.5557464	.1473606	-3.77	0.000	-.8445678	-.2669249
5	.2741977	.110573	2.48	0.013	.0574787	.4909167
year						
0	-.2057627	.1335685	-1.54	0.123	-.467552	.0560267
1	-.2149781	.1369342	-1.57	0.116	-.4833642	.053408
2	-.227502	.1344057	-1.69	0.091	-.4909323	.0359283
3	-.2168636	.1325015	-1.64	0.102	-.4765618	.0428346
4	-.2167557	.1333591	-1.63	0.104	-.4781348	.0446234
5	-.2001909	.1323001	-1.51	0.130	-.4594944	.0591126
6	-.1945682	.1349674	-1.44	0.149	-.4590994	.0699631
7	-.1367099	.1343564	-1.02	0.309	-.4000437	.1266239
8	-.1532571	.1327732	-1.15	0.248	-.4134877	.1069735
9	-.0810621	.1325088	-0.61	0.541	-.3407747	.1786505
10	.002022	.1338693	0.02	0.988	-.260357	.264401
11	-.0197958	.1339719	-0.15	0.883	-.282376	.2427843
12	-.0356323	.1340992	-0.27	0.790	-.2984619	.2271973
13	.0600017	.1334766	0.45	0.653	-.2016076	.3216109
14	.128682	.1304066	0.99	0.324	-.1269101	.3842742
15	.1886627	.1221875	1.54	0.123	-.0508205	.4281459
16	.1267902	.1246797	1.02	0.309	-.1175774	.3711579
17	.0952818	.1260776	0.76	0.450	-.1518257	

```
.3423894
          18  |   .0816317     .12249     0.67   0.505    -.1584444
.3217078
              |
       _cons  |   2.831912    .1686173    16.79   0.000     2.501428
3.162396
-------------+----------------------------------------------------------
------
    /lnalpha  |   -1.61841    .0738158                     -1.763086   -
1.473734
-------------+----------------------------------------------------------
------
       alpha  |   .1982136    .0146313                      .1715147
.2290686
------------------------------------------------------------------------
------
Likelihood-ratio test of alpha=0:   chibar2(01) = 4254.06 Prob>=chibar2 =
0.000

. estat ic

------------------------------------------------------------------------
-----
       Model  |     Obs    ll(null)   ll(model)      df          AIC
BIC
-------------+----------------------------------------------------------
-----
           .  |     490   -2841.038   -2337.806      36     4747.611
4898.61
------------------------------------------------------------------------
-----
               Note:  N=Obs used in calculating BIC; see [R] BIC note.
```

Source: Personal Collection

All coefficients are interpreted relative to a reference age group (likely group 2). Age groups 3–6 and 8 all have significant positive coefficients ($p < 0.001$), meaning they are associated with higher injury rates compared to the reference group. For example, Age group 6 has a coefficient of 0.73, suggesting it has the highest increase in injury rate. Age group 12 has a strong negative coefficient (-1.68, $p < 0.001$), showing substantially lower injury rates than the reference. Age groups 1, 7, 9–11 are not statistically significant, implying no meaningful difference from the reference group in this dataset.

Compared to the reference cause: Cause 2 has a very strong positive effect on injury rates (coefficient = 2.52, $p < 0.001$), suggesting it's a major contributor—likely motor vehicle accidents or another high-impact category. Cause 5 also has a significant positive association with injury rates. Cause 4 shows a significant negative effect, meaning it is associated with lower injury rates. Cause 1 is not significant ($p = 0.711$).

Surprisingly, most year variables are not statistically significant in this model. While coefficients range from negative to slightly positive, none show strong evidence of a clear trend: Years 0–8 generally have small, negative but insignificant coefficients. Year 15 has the highest coefficient (0.189), but it is not significant ($p = 0.123$). This contrasts with the Poisson model, where the year had a significant upward trend—likely due to misfit from overdispersion. In the NBREG model, that trend disappears, suggesting the increase in injury rates over time may not be as strong or consistent when overdispersion is properly handled. The AIC and BIC are more lower than Poisson model which confirms the Negative Binomial model fits the data better (see table 3).

Table 4: Result from Poisson Regression Analysis (using Death rate, Agr group, causes and year)

```
 poisson death_rate b2i.Agegrp b3i.Causes b19.year

Poisson regression                                Number of obs     =
490
                                                  LR chi2(34)       =
462.26
                                                  Prob > chi2       =
0.0000
Log likelihood = -465.97803                       Pseudo R2         =
0.3316

------------------------------------------------------------------------
------
  death_rate |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------+----------------------------------------------------------
------
      Agegrp |
          1  |   1.088204   2.786218     0.39   0.696    -4.372683
6.54909
          3  |    4.09696   2.643979     1.55   0.121    -1.085143
9.279064
          4  |   5.004527   2.639036     1.90   0.058     -.167888
10.17694
          5  |   5.165334   2.638349     1.96   0.050    -.0057351
10.3364
          6  |   5.263678   2.638062     2.00   0.046     .0931707
10.43419
          7  |  -.3677414     4.1894    -0.09   0.930    -8.578814
7.843331
          8  |   4.886201   2.639328     1.85   0.064    -.2867862
10.05919
          9  |   4.376771   2.641805     1.66   0.098    -.8010722
9.554613
         10  |   4.907558   2.639217     1.86   0.063     -.265212
10.08033
         11  |   5.343273   2.637888     2.03   0.043     .1731069
10.51344
         12  |  -.3232732   3.230416    -0.10   0.920    -6.654772
6.008226
             |
      Causes |
          1  |  -2.586948   54.97685    -0.05   0.962    -110.3396
105.1657
          2  |   .0942675   .9241811     0.10   0.919    -1.717094
1.905629
          4  |  -3.451249   1.369964    -2.52   0.012    -6.136329   -
.7661684
          5  |   .4487122   .9241351     0.49   0.627    -1.362559
```

```
2.259984
             |
        year |
          0  |  -1.416621   .3741078    -3.79   0.000    -2.149859   -
.6833835
          1  |  -1.206521   .3575507    -3.37   0.001    -1.907307   -
.5057344
          2  |  -1.138979    .335214    -3.40   0.001    -1.795986   -
.4819715
          3  |  -1.050586   .3143198    -3.34   0.001    -1.666641   -
.4345304
          4  |  -.9755757   .3058195    -3.19   0.001    -1.574971   -
.3761805
          5  |  -.8751583   .2949619    -2.97   0.003    -1.453273   -
.2970436
          6  |  -.8377346   .2997211    -2.80   0.005    -1.425177   -
.2502921
          7  |   -.762669   .2838413    -2.69   0.007    -1.318988   -
.2063503
          8  |  -.7471899   .2821967    -2.65   0.008    -1.300285   -
.1940945
          9  |  -.7124295   .2789003    -2.55   0.011    -1.259064   -
.1657948
         10  |   -.642433   .2726302    -2.36   0.018    -1.176778   -
.1080876
         11  |  -.6296469   .2714901    -2.32   0.020    -1.161758    -
.097536
         12  |  -.5775394   .2668317    -2.16   0.030     -1.10052   -
.0545589
         13  |  -.5152413    .261716    -1.97   0.049    -1.028195   -
.0022874
         14  |  -.4343996   .2552255    -1.70   0.089    -.9346323
.0658331
         15  |  -.3017999   .2446059    -1.23   0.217    -.7812187
.1776189
         16  |  -.2255034   .2397047    -0.94   0.347     -.695316
.2443091
         17  |  -.2428668   .2408978    -1.01   0.313    -.7150179
.2292843
         18  |  -.1956686    .237502    -0.82   0.410     -.661164
.2698269
             |
       _cons |   -4.35515   2.797076    -1.56   0.119    -9.837319
1.127018
-----------------------------------------------------------------------
------

. estat gof

         Deviance goodness-of-fit =   359.959
         Prob > chi2(455)         =    0.9996

         Pearson goodness-of-fit  =  308.9756
         Prob > chi2(455)         =    1.0000

. estat ic

-----------------------------------------------------------------------
-----
```

```
       Model |     Obs   ll(null)   ll(model)      df          AIC
BIC
-------------+---------------------------------------------------------
-----
           . |     490  -697.1094    -465.978      35     1001.956
1148.76
-----------------------------------------------------------------------
-----
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Source: Personal Collection

Each age group is compared to a reference group (likely group 2). Most age groups have positive coefficients, suggesting higher death rates compared to the reference, but only a few are statistically significant: Age group 6 (coef = 5.26, p = 0.046) and Age group 11 (coef = 5.34, p = 0.043) show significantly higher death rates. Groups like 4, 5, 8, and 10 are borderline (p-values between 0.05 and 0.10), indicating weak evidence of increased death rates. Age groups 1, 3, 7, 9, 12 are not significant, suggesting no clear difference from the reference group in those cases.

Cause 4 is the only statistically significant variable (coef = -3.45, p = 0.012), and it's negatively associated with death rate — suggesting this cause leads to lower fatal outcomes. Causes 1, 2, and 5 are not statistically significant, meaning they don't show a strong or consistent effect on death rate in this model.

Most of the year variables are statistically significant and negative, indicating that death rates have generally decreased over time, especially in the earlier years: From 2001 (year 0) through 2013 (year 13), the coefficients are negative and statistically significant ($p < 0.05$), suggesting a steady decline in death rates. After 2013, the coefficients become smaller and non-significant, indicating that the downward trend may have leveled off in recent years.

Pearson $chi^2/df = 308.98 / 455 \approx 0.68$

Deviance $chi^2/df = 359.96 / 455 \approx 0.79$

Both values are less than 1, suggesting underdispersion — a condition where the variance is lower than expected under the Poisson assumption. However, this is less problematic than overdispersion, and model fit remains acceptable (See table 4)

Table 5: Chi-square Analysis (Association between age group and causes of injuries/deaths)

Model	Value	df	p
Pearson Chi-square	211.373	44	<0.0001

From the table below, the Pearson chi-square shows that there is a significant association between the age group and the causes of injuries/deaths since the p-value (<0.0001) < 0.05 (Table 5).This means that the differences in death or injury counts across age groups are not due to chance, and age appears to play a key role in how injuries or deaths are distributed in the population. These findings highlight the importance of considering age-specific factors when designing prevention or intervention strategies.

Table 6: Chi-square Analysis (Association between causes of deaths and number of deaths)

Model	Value	df	p
Pearson Chi-square	1606.554	1492	0.020

From the tables below, the Pearson Chi-square shows that there is asignificant association between the number of deaths and the causes of deaths (cut/pierce, fall, fire/burn, other transport, and poisoning) since the p-value (0.020) < 0.05 (Table 6). This suggests that the number of deaths varies significantly depending on the cause, and the differences observed are unlikely to be due to random chance. Therefore, the cause of death appears to play a meaningful role in the distribution of death cases in the dataset.

Table 7: Chi-square Analysis (Association between causes of injuries and number of injuries)

Model	Value	df	p
Pearson Chi-square	1960	1956	0.470

Similarly, there is no significant association between the number of injuries across the causes of injuries (cut/pierce, fall, fire/burn, other transport, and poisoning) since the p-value (0.0470) > 0.05 (Table 7). This means that, within the data analyzed, the distribution of injuries does not differ significantly across causes. In other words, any differences observed in injury counts among causes of injuries are likely due to random variation rather than a meaningful pattern.

Table 8: Chi-square Analysis (Association between age group and number of deaths)

Model	Value	df	p
Pearson Chi-square	4358.5	4103	0.003

Besides, there is a significant association between the age group across the number of deaths (<1 year, 1-4 years, and so on) since the p-value (0.003) < 0.05 (Table 8). This means that the number of deaths varies meaningfully across different age groups, and the differences observed are unlikely to be due to chance alone. In other words, age appears to be an important factor in the pattern of deaths in this dataset.

Table 9: Chi-square Analysis (Association between age group and number of injuries)

Model	Value	df	p
Pearson Chi-square	5390	5379	0.455

The result shows there is no significant association between the number of injuries across age groups (<1 year, 1-4 years, and so on) since the p-value (0.455) > 0.05 (Table 9). This suggests that, in this dataset, injuries are fairly evenly distributed across age groups, and any differences observed are likely due to chance rather than a meaningful pattern.

Finally, the multiple comparison test is performed to determine the pairs that make a significant contribution and those that do not (see Appendix I).

5 LIMITATIONS AND FUTURE RESEARCH

While this study offers important findings, it also faces certain limitations. The use of retrospective data from a single national source, the CDC, may not capture all contextual variables such as socioeconomic status, geographic variability, or access to healthcare that could influence injury and death patterns. Additionally, the study focused on broad categories like “fall” or “poisoning,” without breaking them down into more specific subtypes or understanding behavioral and environmental influences.

For future research, incorporating more granular data and exploring additional explanatory variables like income levels, education, urban vs. rural settings, and healthcare access could yield even deeper insights. Longitudinal studies examining individual-level trajectories and international comparative analyses would further enrich understanding. Moreover, employing causal inference frameworks may strengthen the interpretation of observed associations and their policy relevance. Future studies could also explore how these patterns have shifted post-2020, particularly in the context of the COVID-19 pandemic and its impact on public health behavior and healthcare access.

6 CONCLUSION

This study set out to examine how age, cause of injury, and time influence injury and death rates in the United States over a 20-year period (2001–2020). Through the use of Poisson and Negative Binomial regression models paired with Chi-square analysis, we uncovered important patterns and associations.

The findings show that injury rates are significantly influenced by both age group and cause of injury, with certain age groups (notably group 6) and causes (like cause 2, possibly motor vehicle-related) consistently associated with higher injury rates. Over time, injury rates showed an increasing trend in the Poisson model; however, once we accounted for overdispersion using a Negative Binomial approach, the year effect became less prominent, suggesting earlier trends may have been overstated. This adjustment strengthens the reliability of the analysis and highlights the importance of selecting appropriate statistical models when dealing with count data.

For death rates, the results revealed a different picture. While age still played a notable role with older age groups showing higher associations with fatal outcomes most causes of death were not significantly different from the reference category. Importantly, the year variable showed a consistent decline in death rates over time, especially in the earlier part of the study period, pointing to possible improvements in healthcare, prevention strategies, or reporting practices.

Taken together, the study highlights the need for age- and cause-specific interventions in public health. Injury prevention efforts should focus on the age groups and causes that carry the greatest risk, while continuing to monitor and adapt strategies as trends evolve over time. These insights not only enhance our understanding of injury and death patterns but also provide a strong foundation for designing evidence-based policies that aim to reduce harm and save lives.

7 REFERENCES

1. Bergen, G., Stevens, M. R., & Burns, E. R. (2016). Falls and Fall Injuries Among Adults Aged ≥65 Years — United States, 2014. MMWR. Morbidity and Mortality Weekly Report, 65(37), 993–998. https://doi.org/10.15585/mmwr.mm6537a2
2. Moreland, B. L., Burns, E. R., & Haddad, Y. K. (2022). Differences in fall-related emergency departments visits with and without an Injury, 2018. Journal of Safety Research, 82, 367–370. https://doi.org/10.1016/j.jsr.2022.07.002
3. Rajabali, F., Turcotte, K., Zheng, A., Purssell, R., Buxton, J. A., & Pike, I. (2023). The impact of poisoning in British Columbia: a cost analysis. CMAJ Open, 11(1), E160–E168. https://doi.org/10.9778/cmajo.20220089
4. Shiue, K. Y., Naumann, R. B., Proescholdbell, S., Cox, M. E., Aurelius, M., & Austin, A. E. (2022). Differences in overdose deaths by intent: Unintentional & suicide drug poisonings in North Carolina, 2015–2019. Preventive Medicine, 163. https://doi.org/10.1016/j.ypmed.2022.107217
5. WHO. (n.d.). Global burn registry. Retrieved July 9, 2023, from https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/burns/global-burn-registry
6. WHO. (2022). Road traffic injuries. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries
7. Potoglou, D., Carlucci, F., Cirà, A., & Restaino, M. (2018). Factors associated with urban non-fatal road-accident severity. International Journal of Injury Control and Safety Promotion, 25(3), 303–310. https://doi.org/10.1080/17457300.2018.1431945
8. Wang, J., Lu, H., Sun, Z., Wang, T., & Wang, K. (2020). Investigating the impact of various risk factors on victims of traffic accidents. Sustainability (Switzerland), 12(9). https://doi.org/10.3390/su12093934
9. Panda, C., Mishra, A. K., Dash, A. K., & Nawab, H. (2022). Predicting and explaining severity of road accident using artificial intelligence techniques, SHAP and feature analysis. International Journal of Crashworthiness. https://doi.org/10.1080/13588265.2022.2074643
10. Santos, K., Firme, B., Dias, J. P., & Amado, C. (2023). Analysis of Motorcycle Accident Injury Severity and Performance Comparison of Machine Learning Algorithms. Transportation Research Record. https://doi.org/10.1177/03611981231172507
11. Chen, S., Kuhn, M., Prettner, K., & Bloom, D. E. (2019). The global

macroeconomic burden of road injuries: estimates and projections for 166 countries. The Lancet Planetary Health, 3(9), e390–e398. https://doi.org/10.1016/S2542-5196(19)30170-6
12. Gorucu, S., Michael, J., & Chege, K. (2022). Nonfatal Agricultural Injuries Treated in Emergency Departments: 2015-2019. Journal of Agromedicine, 27(1), 41–50. https://doi.org/10.1080/1059924X.2021.1913271
13. Khurana, B., Hines, D. A., Johnson, B. A., Bates, E. A., Graham-Kevan, N., & Loder, R. T. (2022). Injury patterns and associated demographics of intimate partner violence in men presenting to U.S. emergency departments. Aggressive Behavior, 48(3), 298–308. https://doi.org/10.1002/ab.22007
14. WHO. (2021). Falls. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/falls
15. Lytras, D., Sykaras, E., Iakovidis, P., Kasimis, K., Myrogiannis, I., & Kottaras, A. (2022). Recording of Falls in Elderly Fallers in Northern Greece and Evaluation of Aging Health-Related Factors and Environmental Safety Associated with Falls: A Cross-Sectional Study. Occupational Therapy International, 2022. https://doi.org/10.1155/2022/9292673
16. Zhang, K., Qi, J., Zuo, P., Yin, P., Liu, Y., Liu, J., Wang, L., & Li, L. (2022). The mortality trends of falls among the elderly adults in the mainland of China, 2013—2020: A population-based study through the National Disease Surveillance Points system. The Lancet Regional Health - Western Pacific, 19. https://doi.org/10.1016/j.lanwpc.2021.100336
17. Li, X., Zhuge, C., & Yu, B. (2019, August). Analysis on the impact of illegal driver behaviors on road traffic accidents case study on China. In 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC) (Vol. 1, pp. 250-253). IEEE.
18. Lerdsuwansri, R., Phonsrirat, C., Prawalwanna, P., Wongsai, N., Wongsai, S., & Simmachan, T. (2022). Road traffic injuries in Thailand and their associated factors using Conway-Maxwell-Poisson regression model. Thai Journal of Mathematics, 240-249.
19. Rathnayake, C. W., Jones, S., Soto-Berelov, M., & Wallace, L. (2022). Human–elephant conflict and land cover change in Sri Lanka. Applied geography, 143, 102685.
20. Neupane, D., Johnson, R. L., & Risch, T. S. (2013, August). Temporal and spatial patterns of human-elephant conflict in Nepal. In 2013 international elephant & rhino conservation & research symposium proceedings (pp. 1-11).
21. Koutras, K., Kasfiki, E., & Pyrgioti, E. (2021, September). Statistical Analysis of Deaths and Damages Caused by Lightning Strikes in Greece. In 2021 35th International Conference on Lightning Protection (ICLP)

and XVI International Symposium on Lightning Protection (SIPDA) (Vol. 1, pp. 01-05). IEEE.

22. Turner, J. C., Leno, E. V., & Keller, A. (2013). Causes of mortality among American college students: A pilot study. Journal of college student psychotherapy, 27(1), 31-42.
23. Halabi, Y., Xu, H., Long, D., Chen, Y., Yu, Z., Alhaek, F., & Alhaddad, W. (2022). Causal factors and risk assessment of fall accidents in the US construction industry: A comprehensive data analysis (2000–2020). Safety science, 146, 105537.
24. Malin, F., Silla, A., & Mladenović, M. N. (2020). Prevalence and factors associated with pedestrian fatalities and serious injuries: case Finland. European transport research review, 12(1), 1-17.
25. Wang, H., Liu, Z., Wang, X., Graham, T., & Wang, J. (2021). An analysis of factors affecting the severity of marine accidents. Reliability Engineering & System Safety, 210, 107513.
26. Taylor, M., Francis, H., & Fielding, J. (2023). Old age and fire injury. Journal of Fire Sciences, 41(1-2), 16-31.

8 AUTHORS

Toluwalase Janet Akingbade currently serves as a lecturer in the Department of Industrial Mathematics at Covenant University, Ota, Nigeria. In this role, she imparts knowledge of Statistics to both undergraduate and postgraduate students. With nearly a decade of experience in lecturing Statistics, Dr. Akingbade earned her B.Sc. (Hons) and M.Sc. in Statistics from the University of Ilorin, Nigeria, and holds a Ph.D. from the University of Nigeria, Nsukka, where her research focused on Survey Methodology.
Actively engaged in the academic community, Dr. Akingbade is a member of the Nigerian Statistical Association. She has contributed to reputable journals through published articles and has actively participated in conferences. Dr. Akingbade's research interests encompass Sample Surveys, Biostatistics, and Machine Learning.

Oluwafemi Samson Balogun is a former lecturer at the Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Adamawa State, Nigeria. He is an International Visiting Scholar at the Department of Mathematics, Covenant University, Ota, Ogun State, Nigeria and he is currently a post-doctoral researcher at the School of Computing, University of Eastern Finland. He holds a Ph.D. in Statistics from the Department of Statistics, University of Ilorin, Kwara State, Nigeria. He is a senior research fellow at the Centre for Multidisciplinary Research and Innovation (CEMRI) and He is a member of the International Biometric Society (IBS), International Statistical Institute (ISI), Nigerian Statistical Association (NSA), Nigerian Mathematical Society (NMS). He has several conference and journal papers in reputable international bodies, and he has also served as a guest reviewer for several journals. His research interests are data science, machine learning, data mining, biostatistics, categorical data analysis, modeling, and probability distribution models.

Donald Douglas Atsa'am is a Senior Lecturer of Data Science in the Department of Computer Science, Joseph Sarwuan Tarka University Makurdi, Nigeria. He holds a PhD in Applied Mathematics and Computer Science from the Eastern Mediterranean University, Famagusta, North Cyprus. His research interests are in Data Science, Machine Learning, and Knowledge Discovery.

Samuel Nii Odoi Devine is a Presbyterian University, Ghana, ICT faculty member. He holds an MSc in Information Technology and a Ph.D. Computer Science candidate at the University of Johannesburg. In 2019, his paper was nominated with two others for the best paper award at the knowledge management and information systems conference held in Vienna, Austria. He has several years of teaching experience and has been involved in several IS projects. His research interest is multidisciplinary, spanning Machine Learning Applications, Natural Language Processing, Big Data, and

Knowledge Management.

Temidayo O. Omotehinwa, Ph.D., received his B.Sc., M.Sc., and Ph.D. in Computer Science from the University of Ilorin, Nigeria, in 2007, 2011, and 2017 respectively. He currently holds the position of Senior Lecturer and acting head of the Department of Mathematics and Computer Science at the Federal University of Health Sciences, Otukpo, Nigeria. Dr. Omotehinwa specializes in machine learning, knowledge extraction, and data analytics, with notable contributions to the field through his published work in prestigious journals and conferences. In 2023, he was awarded an all-expense paid grant by Luxembourg Aid & Development via INSTITUT PASTEUR de DAKAR. This grant allowed him to bring his expertise to the first International Course on Genomic Algebra Theory, Tools, and Applications (GATTA) in Dakar, Senegal.

Appendix 1

Variable Categories

Age group	Categories	Causes	Categories	Year	Categories
<1 year	1	Cut/Pierce	1	2001	0
1-4 years	2	Fall	2	2002	1
5-9 years	3	Fire/burn	3	2003	2
10-14 years	4	Other transport	4	2004	3
15-24 years	5	Poisoning	5	2005	4
25-34 years	6			2006	5
35-44 years	7			2007	6
45-54 years	8			2008	7
55-64 years	9			2009	8
65-74 years	10			2010	9
75-84 years	11			2011	10
85+ years	12			2012	11
				2013	12
				2014	13
				2015	14
				2016	15
				2017	16
				2018	17
				2019	18
				2020	19