Final Report
WBS: 25-1121-3002-103
Report # MATC: 103

Enhancing Safety and Well Being of Aging Workforce within the Transportation Sector

An Association Rule Data Mining Approach for Understanding Causes and Impacts of Worker Compensation Claims

Islam H. El-adaway, Ph.D.
Associate Dean for Academic Partnerships
Hurst-McCarthy Professor of Construction Engineering and Management
Professor of Civil Engineering
Founding Director of Missouri Consortium for Construction Innovation
Department of Civil, Architectural, and Environmental Engineering
Department of Engineering Management and Systems Engineering
Missouri University of Science and Technology

Mariam Elazhary
Ph.D. Student
Department of Civil, Architectural, and Environmental Engineering
Missouri University of Science and Technology

2025

A Cooperative Research Project sponsored by the U.S. Department of Transportation-Office of the Assistant Secretary for Research and Technology.

Enhancing Safety and Well Being of Aging Workforce within the Transportation Sector: An Association Rule Data Mining Approach for Understanding Causes and Impacts of Worker Compensation Claims

Islam H. El-adaway, Ph.D
Associate Dean for Academic Partnerships
Hurst-McCarthy Professor of Construction Engineering and Management
Professor of Civil Engineering
Founding Director of Missouri Consortium for Construction Innovation
Department of Civil, Architectural, and Environmental Engineering
Department of Engineering Management and Systems Engineering
Missouri University of Science and Technology

Mariam Elazhary
Ph.D. Student
Department of Civil, Architectural, and Environmental Engineering
Missouri University of Science and Technology

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska–Lincoln

May 2025

Technical Report Documentation Page

<table>
<tr><td colspan="2">1. Report No.
25-1121-3002-103</td><td colspan="2">2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="4" rowspan="2">4. Title and Subtitle
Enhancing Safety and Well Being of Aging Workforce within the Transportation Sector: An Association Rule Data Mining Approach for Understanding Causes and Impacts of Worker Compensation Claims</td><td colspan="2">5. Report Date
May 2025</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="4">7. Author(s)
Islam El-awaday, ORCID No. 0000-0002-7306-6380</td><td colspan="2">8. Performing Organization Report No.
25-1121-3002-103</td></tr>
<tr><td colspan="4" rowspan="2">9. Performing Organization Name and Address
Mid-America Transportation Center
Prem S. Paul Research Center at Whittier School
2200 Vine St.
Lincoln, NE 68583-0851</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
69A3552348307</td></tr>
<tr><td colspan="4" rowspan="2">12. Sponsoring Agency Name and Address
Office of the Assistant Secretary for Research and Technology
1200 New Jersey Ave., SE
Washington, D.C. 20590</td><td colspan="2">13. Type of Report and Period Covered
Final Report
June 2023 – May 2025</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="6">15. Supplementary Notes</td></tr>
<tr><td colspan="6">16. Abstract
Safety within the transportation sector has long been a concern for researchers and industry stakeholders. High injury rates in the transportation sector and an increase in expenditure toward transportation projects, state departments of transportation put in place key performance indicators aiming for a safer workforce. To do so, root causes of safety incidents must be studied to help prevent/mitigate future incidents. Previous studies addressed this using various analytical approaches; however, they have either failed to account for both internal and external factors, only accounted for one factor at a time affecting safety, or conducted a methodology requiring extensive, confidential data that is difficult to retrieve. The goal is to decipher combinations of internal and external reasons behind occupational injuries in the transportation sector by utilizing top-level data without the need to access detailed confidential records. This was accomplished through a systematic methodological approach, employing a funneled multiple linear regression (MLR), that offered a novel way of evaluating the influence of different variable combinations on safety outcomes. The methodology was composed of (1) identifying independent and dependent variables from publicly available data by the Bureau of Labor Statistics, (2) studying multicollinearity, (3) conducting feature selection, and (4) applying a funneled MLR approach to examine all possible variable combinations. The findings included 5 high-performing MLR models composed of critical combinations of variables impacting safety incidents in the transportation sector. The contribution of this research is providing actionable insights for developing targeted interventions aimed at mitigating safety risks and enhancing workplace safety within the transportation sector.</td></tr>
<tr><td colspan="3">17. Key Words
Safety, Transportation Sector, Enhancing Workforce Retainage</td><td colspan="3">18. Distribution Statement</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td colspan="2">20. Security Classif. (of this page)
Unclassified</td><td colspan="2">21. No. of Pages
62</td><td>22. Price</td></tr>
</table>

Table of Contents

List of Figures

List of Tables

Acknowledgments

The U.S. Department of Transportation provided financial support for this research through the University of Nebraska-Lincoln Mid-America Transportation Center under contract number 69A3552348307.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Abstract

Safety within the transportation sector has long been a concern for researchers and industry stakeholders. With high injury rates in the transportation sector and an increase in expenditure toward transportation projects, state departments of transportation put in place key performance indicators aiming for a safer workforce. To do so, root causes of safety incidents must be studied to help prevent/mitigate future incidents. Previous studies addressed this using various analytical approaches; however, they have either failed to account for both internal and external factors affecting safety, or conducted a methodology requiring extensive, confidential data that is difficult to retrieve. The goal is to decipher combinations of internal and external reasons behind occupational injuries in the transportation sector by utilizing top-level data without the need to access detailed confidential records. This was accomplished through a systematic methodological approach, employing a funneled multiple linear regression (MLR), that offered a novel way of evaluating the influence of different variable combinations on safety outcomes. The methodology was composed of (1) identifying independent and dependent variables from publicly available data by the Bureau of Labor Statistics, (2) studying multicollinearity, (3) conducting feature selection, and (4) applying a funneled MLR approach to examine all possible variable combinations. The findings included five high-performing MLR models composed of critical combinations of variables impacting safety incidents in the transportation sector. This research contributes actionable insights for developing targeted interventions aimed at mitigating safety risks and enhancing workplace safety within the transportation sector.

Chapter 1 Introduction

The construction industry is one of the most hazardous industries in the world (Chiang et al., 2018). An important segment of the industry is the transportation sector. The need to build safe and efficient transportation projects is an essential part of an economy's growth and health. This can be seen in the new Infrastructure Investment and Jobs Act (IIJA) signed by President Biden in November 2021 with a $1.2 trillion budget for transportation and infrastructure (USDOT, 2023a). To maintain the resilience of the transportation sector, recruiting and retaining a highly skilled workforce is key (Metro et al., 2021). Following such, the ultimate goal would be to increase the well-being and productivity of the retained workforce to achieve project success (Karthick et al., 2021). A vital aspect to account for is the safety of the workforce since projects rely heavily on human power (Ayhan & Tokdemir, 2020a). This is necessary due to the organizational and technical issues that safety incidents cause (Hadikusumo et al., 2017). The transportation sector is no different. In fact, one of the strategic objectives of the U.S. Department of Transportation (USDOT) 2022-2024 Performance Plan, is to maintain the "*health safety and well-being"* of transportation workers (USDOT, 2023b). One of the key performance indicators set in the 2022-2024 plan, is to decrease the transportation worker injury rate by 2026.

Analyzing occupational injury data is necessary to be proactive in eliminating the reason(s) behind injuries. Marji et al. (2023) have recommended consistently analyzing injury data to avoid hazards. Following the same logic, understanding how safety accidents occur has been extensively advocated for in the literature (Assaad & El-adaway, 2021a; Ayhan & Tokdemir, 2020a; Gibb et al., 2014; Y. Li & Bai, 2008). While singling out primary factors contributing to injuries offers a swift and convenient approach to analysis, it may not offer a comprehensive perspective necessary for effectively addressing and mitigating hazards.

Meaning, that the causes for safety incidents must be looked at as a combination of contributing factors since injuries in the construction sector are complex and multi-causal (Suraji et al., 2001). To retrieve such results, an ideal level of detail for data must be available to analyze. However, due to the confidential nature of the data and as per the U.S. Privacy Act of 1974, the full data disclosure may not be easily accessible for detailed analysis. For example, one of the most renowned sources of injury cases is the BLS Injuries, Illnesses, and Fatalities (IIF) program, but it only provides top-level rates or counts (Choe & Leite, 2017).

Existing studies employ various analytical methods to retrieve causes of safety incidents in the transportation sector, but their scope and depth often prove short of something critical. A fundamental obstacle arises due to the limited availability of safety data, often attributed to security concerns, which impedes a thorough understanding. Despite scholars accessing data and presenting insightful discoveries, many studies primarily concentrate on individual safety incidents, disregarding the intricate interplay of various factors and the importance of examining interdependencies (Koulinas et al., 2023). Additionally, studies with unrestricted access to safety data frequently overlook both internal and external safety concerns, as seen in research on work zone crashes focusing solely on traffic accidents without considering other onsite risks such as falls (Das et al., 2023). Even studies attempting to circumvent this issue by gathering data through surveys (Marji et al., 2023) still encounter limitations due to differences in perception and memory constraints (Ghosh et al., 2023).

Based on the presented gap, the goal of this research is to decipher combinations of internal and external reasons behind occupational injuries in the transportation sector by utilizing top-level data without the need to attain access to detailed confidential records. The goal shall be accomplished by identifying internal and external sources of safety issues in the transportation sector and establishing a multi-tier methodology that can provide comprehensive deductions. Such deductions shall proactively contribute to prolonged workforce retainage and enhance the industry's image. Further, it promotes practicality and facilitates the implementation of findings in safety data.

Chapter 2 Background Information

2.1 Construction Sector Safety-Related Studies

In a general sense, construction-related safety incidents have been extensively approached throughout the literature. Some studies achieve a proactive approach by trying to identify and avoid factors that cause safety accidents. For example, Ayhan & Tokdemir (2020) have created a model to predict outcomes of construction safety accidents and then provided recommendations for avoidance. Fass et al. (2017) focused on fall and struck-by accidents by analyzing 519 incident reports in the Arabian Gulf region to explore the key factors affecting safety incidents and how to avoid them. Moreover, Assaad & El-adaway (2021) examined 100 fatal accident case files to determine the most critical combination of causes for fatal accidents and to avoid them. All of the above studies utilized different approaches, but their goal was to identify and eliminate root causes of safety accidents in construction. Another aspect of risk management is mitigating safety incidents when they cannot be avoided due to the nature of construction projects. Mohandes et al. (2022) focused on the risk aspect of safety and aimed to enhance how construction projects may better deal with and mitigate safety-related risks. While Muzafar (2021) has opted to include building information technology (BIM) to accurately identify safety-related risks on site.

A more in-depth approach to identifying causes of safety issues is exploring human cognition and how it affects safety behavior on site. Liu et al. (2023) conducted a review and discussion about the various antecedents that affect a worker's cognition during safety issues and recommended that construction managers utilize the understanding of such antecedents to reduce unsafe behavior. Hu et al. (2023) have corroborated the aforementioned research through a state-of-the-art review and emphasized how safety interventions must be personalized for workers through understanding their cognitive status. Discussion in the literature has taken cognition a step further by using an electroencephalogram (EEG) wearable device to capture the brain activity of workers responding to hazards (Zhou & Liao, 2023). The authors were capable of relying on human-machine collaboration to formulate a tool that predicts the cognitive response of workers to hazards. Using the same principle, various studies have utilized EEG to study the cognitive behavior of workers to prevent, predict, or mitigate safety issues (Jeon & Cai, 2022, 2023; Mehmood et al., 2023).

Other researchers have chosen to focus on the imperative aspect of safety imposed by organizations through safety training. Safety training was a significant development in the construction industry in recognizing safety hazards and safety-related risks (Namian et al., 2016). However, safety training may not always be convey safety as a big issue for construction workers (Loosemore & Malouf, 2019). Fu et al. (2024) used eye-tracking technology to see the effect of different cues on effective safety training. Meanwhile, Gao et al. (2019) have concluded that computer-aided safety training is more effective than traditional methods. That is why computer-aided training was taken to another level by introducing Industry 4.0, and including virtual reality to create an enhanced and engaging safety training experience (Bader et al., 2024; Dang et al., 2024; Rokooei et al., 2023).

As seen above, the discussion of safety in the construction realm is vast due to its significance; indeed, studies expand beyond the areas discussed. For example, some studies choose to focus on migrant workers due to their high susceptibility to safety accidents (Al-Bayati, Eiris, et al., 2023; Lyu et al., 2023; Nielsen et al., 2023). Other studies revolve around resolving a certain type of incident such as falling from heights (Tözer et al., 2024; Zermane et al., 2023), struck-by incidents (Bobadilla et al., 2014; H. Kim et al., 2023), fatal incidents (Park et al., 2020; Zermane et al., 2023), etc. Meanwhile, researchers may opt to look at the bigger picture by studying specific types of projects that are deemed to be relatively more hazardous. Some examples include high-rise buildings (Manzoor et al., 2021), tunnels and bridges (Spangenberg et al., 2005; Ye et al., 2023), and highways (Nnaji et al., 2020). While the discussion of safety in the construction sector may contain various divisions, for the purpose of this study, the focus shall be projects in the transportation sector.

2.2 Transportation Sector Safety-Related Studies

The topic of safety in the transportation sector is one of great interest and importance due to its hazardous conditions compared to other sectors. The transportation sector workforce is constantly exposed to passing traffic, extreme conditions, and heavy moving equipment (Al-Shabbani et al., 2018). This kind of exposure may result in injuries or even fatalities within the workplace. The amount of occupational injuries and illnesses per total number of employees in the transportation sector is 1.5 times more than in the buildings sector (BLS, 2023b); thus, focusing on occupational injuries occurring amongst the transportation sector workforce is justified. With the increase of governmental spending in the transportation sector reaching a projected 2024 budget of $145 billion, (U.S. Department of Transportation, 2024), safety incidents are more inclined to increase. The statistic was corroborated by Harris et al. (2022) through a simple frequency analysis of safety incidents in the transportation sector using data from the Bureau of Labor Statistics (BLS) and the National Highway Traffic Safety Administration (NHTSA). The alarming trends in injuries have created the need to understand and analyze the causes of safety incidents to protect the transportation workforce.

Three types of analysis can be commonly seen in the transportation sector safety-related studies: descriptive, predictive, and prescriptive (Cote, 2021). Descriptive analytics involves identifying trends and describing what happened through association algorithms or clustering analysis; while predictive, allows for building models that can predict future events from historical data such as using artificial neural networks (Williams, 2011). The prescriptive analysis advises on how to move forward (Balali et al., 2020). The prescriptive analysis method can incorporate descriptive and predictive methods and serve as grounds for further discussion or be used on its own. This section offers a detailed discussion on literature addressing all four data analytic methods.

A simple approach to descriptive analysis was undertaken by Al-Bayati et al. (2023). The authors attained 75 reports of fatal construction incidents reported by the Fatality Assessment and Control Evaluation (FACE) Program. Descriptive content analysis was done to identify and rank the external and internal contributors during road works. The authors later proceeded to provide prescriptive recommendations. However, it must be noted that this simple technique has not allowed for identifying possible combinations of reasons that might have led to the safety incidents.

Several research efforts utilized a descriptive method called the association rule algorithm to find a combination of factors that cause work zone crashes (Chammout et al., 2024; Das et al., 2023; Weng et al., 2016). These researchers utilized data related to work zone crashes that were readily available at sources such as the *Michigan work zone crash* data, the *Fatality Analysis Reporting System*, the *US National Highway Traffic Safety Administration, and* others. The data is usually exhaustive enough to allow for association rules to be conducted in a meaningful manner. Valcamonico et al. (2022) provided another novel approach to describing work zone crashes by using natural language processing (NLP) to classify accidents. Nevertheless, investigating work zone crashes solely delves into the various situations that result in work zone traffic accidents, which is crucial but does not offer a complete perspective. These investigations often overlook internal factors occurring on the site related to construction activities, such as falls from heights or injuries caused by explosives. Other attempts have been made in descriptive analytics when data was not available and Delphi questionaries were utilized (Hallowell et al., 2011). However, a prevailing issue in questionnaires is the reliability of the data. Yet, some researchers have had access to exhaustive data on more safety incidents in transportation construction sites than just work zone crashes. An example of this is a study conducted by Y. A. Kim et al. (2013) who analyzed accidents in Korea using analysis of variance and cross-tabulation to retrieve the effect of different factors on highway safety incidents. This study utilized cross-tabulation to retrieve the relative importance of factors that can cause a safety issue. All observed factors were related to the surroundings, and the characteristics of the workers were not accounted for, which could be highly influential (Lyu et al., 2023). Another study on railway construction demonstrated the opposite by focusing only on worker behavior and not the surroundings (Guo et al., 2021). Therefore, exploring both the behavior and worker characteristics that led to an accident as well as the surroundings, in other words the internal and external factors, is important. Research done by Tong et al. (2020) has also utilized an association rule algorithm to retrieve a combination of causes leading to highway construction accidents in China while taking into account external and internal factors. In this study, data accessibility was not an issue, allowing for a detailed analysis, and the focus was targeted towards highway construction without accounting for other major project types in the transportation sector.

The second most popular method in the literature is the prescriptive method. Under this method, authors have observed specific safety issues and proposed solutions through a framework, a decision support system, new tools, technology integration, etc. Esmaeili & Hallowell (2013) incorporated common safety risks into a decision support system to enhance the integration of the safety risks into the project schedule. However, no attention was provided to providing enhancement measures for the workforce. On the other hand, Nnaji et al. (2018) proposed a decision-making system to select the optimal safety technologies for highway construction, with a focus directed towards transportation incidents. Due to the hazardous nature of transportation projects, it is important to consider all aspects when redesigning safety training for the highway construction workforce (Ammar & Dadi, 2023). While safety training often accounts for factors under control of the workforce, safety incidents may occur that include mishappenings out of the workers' control. For example, poor visibility at night is an inevitable aspect of the job, which is why Nnaji et al. (2020) focused on worker visibility at nighttime. The authors proposed and evaluated the effects of wearable lighting systems as a solution. On the same topic of visibility, Arditi et al. (2004) explored and recommended the best safety vests concerning luminance. While focusing on critical aspects to solve is commendable in a practical sense, there still exists a need to explore what combination of happenings would exacerbate the occurrence of safety incidents. Bad nighttime visibility may exacerbate safety incidents even more when coupled with a certain occupation or a certain gender. Therefore, this may alter the prescriptive solution provided. More recent sophisticated approaches attempting to solve this integrated technologies such as augmented reality and digital twins. Sabeti et al. (2021) utilized augmented reality to create a real-time notification system to notify highway construction workers of any upcoming hazards. Ye et al. (2023) diverted attention to tunnel projects and used digital twins to virtually simulate early hazard warnings and provide solutions through emergency response plans. However, the above solutions rely on predefined hazards recognized individually such as upcoming car accidents rocks or falling in tunnels. To provide more heuristic warning signs there must be rigorous acknowledgment of critical factors causing them.

Lastly, predictive analytics, while prominent under the study of safety in the construction sector (Ayhan & Tokdemir, 2020a; Bobadilla et al., 2014; Ghodrati et al., 2018; Zermane et al., 2023), tailoring it to the transportation sector is an emerging area (Bortey et al., 2022). Alqatawna et al. (2021) and Li & Yu (2021) created models to predict highway traffic accidents but did not address any internal construction-related issues. Predictive models can be beneficial in becoming proactive but would require extensive and detailed data sets. Regardless of the analytic method utilized, all authors had the common aim of preventing or mitigating future safety issues.

The literature on safety in the transportation sector exposes significant gaps in addressing comprehensive safety concerns. While prevailing studies utilize three common analytical approaches—descriptive, predictive, and prescriptive—their coverage and depth often fall short. A primary challenge emerges regarding security limiting the availability of safety data that hinders a complete understanding of transportation incidents. Despite scholars accessing data and presenting insightful findings, inherent limitations compromise the overall comprehensiveness of results. Notably, many studies focus on individual safety incidents, neglecting the intricate interplay of multiple factors. For instance, the significance of two non-critical reasons occurring simultaneously might surpass that of a top-ranked issue alone; thus, examining interdependencies is crucial (Koulinas et al., 2023). Furthermore, studies with unrestricted safety data often overlook both internal and external safety issues, as evident in research on work zone crashes that concentrate solely on traffic accidents without considering other on-site risks like falls. Other studies tried to evade the problem by collecting data through surveys (Marji et al., 2023), though they still faced the limitation of perception differences and memory restrictions (Ghosh et al., 2023). Moreover, a prevailing issue is the narrow focus on specific project types within the transportation sector, overlooking the broader context.

Chapter 3 Methodology

This section introduces a systematic approach for investigating factors contributing to nonfatal safety incidents in transportation using Multiple Linear Regression (MLR). By utilizing data from the BLS's SOII and integrating internal and external variables impacting safety outcomes, the methodology aims to identify high-performing models with various combinations of variables. Unlike traditional approaches seeking a single optimal MLR model, this method explores multiple models to understand safety dynamics better. It proposes a novel methodology for systematically evaluating the impact of different variable combinations on safety using MLR, maintaining statistical thresholds. This approach offers an alternative to association rules, allowing exploration of influential factor combinations without the need for individual datasets. Figure 3.1 provides a simple, heuristic demonstration of the four stages. Stage 1 is identifying the independent and dependent variables for the MLR. Stage 2 is studying the correlation between the independent variables to retrieve the number of variables to use: feature size. Stage 3 is conducting two types of feature selection methods to rank the variables based on their influence on the independent variable. Finally, stage 4 consists of evaluating the independent variables using MLR through a funneling effect. Statistical analysis in stages 1, 2, and 4 was conducted using JASP 0.18.1 software, while stage 3 was conducted using Python.

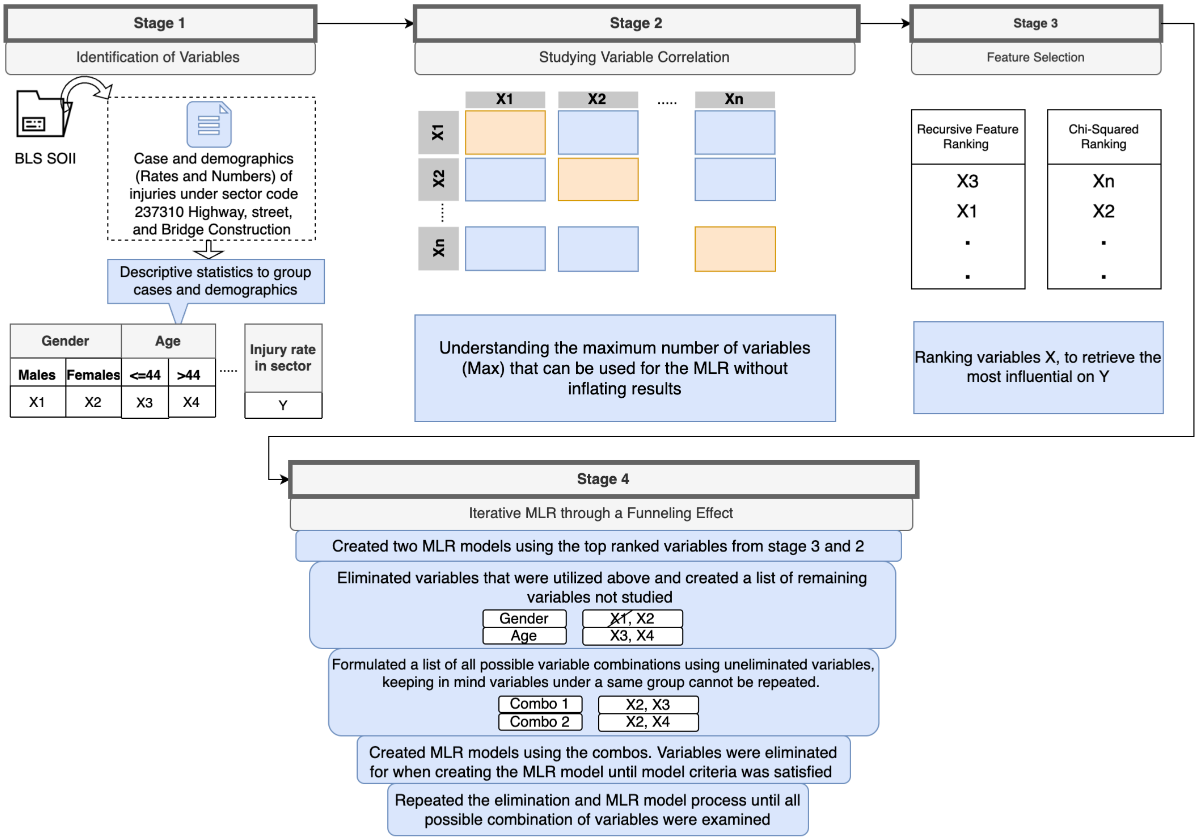

Figure 3.1 Research Methodology

3.1 Stage 1: Identification of Variables

To establish the variables, the BLS SOII data was retrieved for the case and demographics injuries under the transportation sector were identified as 237310 Highway, Street, and Bridge construction. The SOII is a survey conducted by the BLS to gather work-related injuries or illnesses that go beyond first aid treatment (BLS, 2023c). The said data has been used to observe trends, patterns, and frequency by various literature (Dong et al., 2013; Harris et al., 2022); however, as mentioned previously, this research has utilized the data using a different approach. Initially, two reports were generated that encompassed the count of injuries and the rate of injuries. The former was used for the independent variables while the latter was used for the dependent variable.

The number of illnesses and demographics reported by the BLS are available from 2011 to 2020 and are divided into different cases and demographics. This research has accounted for nine major relevant divisions that are composed of external and internal factors: gender, age, occupation, length of service with employer, race, event/exposure, day of the week, time of the day, and hours worked. Even though the BLS reports six other divisions, they were not considered by the authors since they do not account for aspects present before the incidents or features and are instead classified as the aftermath of the injury. An example of such would be the "*nature of injury*" or the "*part of body affected*"*;* these divisions would not explain reasons behind safety incidents. Under each major division, there are subdivisions; for example, the length of service division is divided into less than three months, three months to 11 months, one year to five years, and more than five years. However, this would mean the number of independent variables would be too much and difficult to manage; therefore, to be efficient and retrieve meaningful results, the subdivisions were grouped as in Alexander et al. (2017). Sanni et al. (2021) stressed the importance of avoiding the asymmetric distribution of groups; therefore, descriptive statistics were used to group subdivisions according to the median. Accordingly, the median was chosen to group subdivisions according to the central tendency of the values (Cooksey, 2020). Figure 3.2 demonstrates an example of the conducted grouping where the accumulative sum of the subadvisors was plotted, and the median formulated the breaking point where groups were made. This was done for all divisions to retrieve a final list of independent variables (count of illnesses and injuries for every grouped subdivision) under various divisions. Every variable would attain 10 data points from 2011 to 2020.

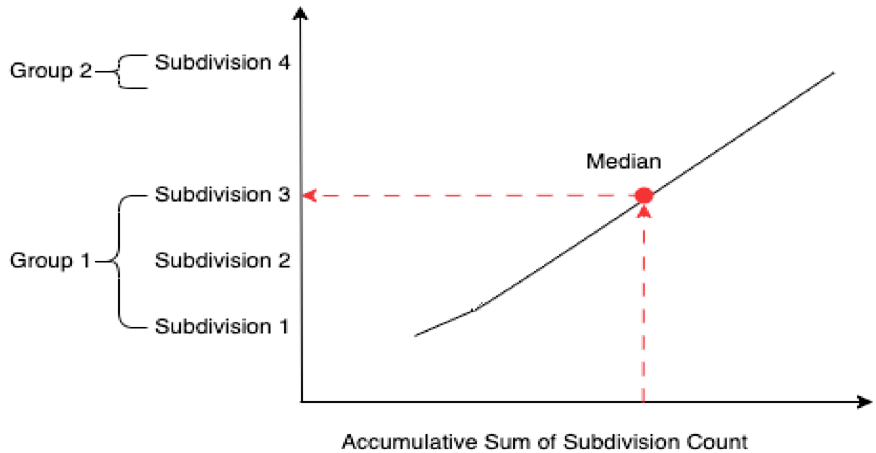

Figure 3.2 Grouping Subdivisions of Independent Variables

The second report is the rate of injuries and illnesses in the transportation sector. It is the number of injuries and illnesses per 100 full-time workers and calculated using the number of injuries and illnesses over the total hours worked by all employees during the calendar year multiplied by 200,000, which is the base for 100 equivalent full-time workers (BLS, 2023a). Utilizing safety rates as independent variables is a strategy utilized by various researchers (Alexander et al., 2017; Ghodrati et al., 2018). The rates were also obtained from the available years, 2011-2020, and were used as the dependent variable. Accordingly, dependent and independent variables would have 10 data points to conduct an MLR analysis; this number of data points is sufficient as per the 10-15 established range of minimum points stated by Harrell (2001) and Ghodrati et al. (2018). However, it must be noted that the SOII does attain a major limitation as critiqued and discussed by the (BLS, 2023c) in that there is evidence of a possible undercount. This could be due to poor record-keeping by employers, failure to capture injuries and illnesses with long onset, as well as excluding self-employed and private household workers.

However, since an undercount might affect both the dependent and independent variables similarly, the limitation would not have a major effect on the validity of the results yielded.

3.2 Stage 2: Studying Variable Correlation

The second stage involves studying the multicollinearity between the independent variables through Pearson's correlation test. This is a nonparametric test that yields numbers between -1 and 1 where the higher the absolute value the stronger the linear relation. Construction management literature addressing MLR has highlighted that the presence of highly correlated independent variables would lead to an inflation of the model's performance (Y. Zhang et al., 2017). Therefore, dealing with multicollinearity is a critical prerequisite to avoid reporting a false combination of factors that affect the rate of safety incidents in the transportation sector. Two common approaches to deal with multicollinearity are Least Absolute Shrinkage and Selection Operator (LASSO) regression and Variance Inflation Factor (VIF). Tong et al. (2021) have utilized LASSO to deal with inflation and defined it as a method suitable for feature selection that enhances the prediction capability of the model when independent variables have high multicollinearity. However, as previously mentioned, the goal of this research is to examine the highly influential combination(s) of variables on the rate of safety incidents in the transportation sector and not to predict the rate. Therefore, an approach similar to Arora et al. (2023) and Garg & Misra (2021) was utilized where the high multicollinearity results from the Pearson correlation test in this stage were further dealt with by using VIF to eliminate variables that may inflate results when constructing the model. The VIF approach shall be further explained in stage 4 describing the application of MLR. Moreover, this stage also resulted in retrieving the maximum number of variables used to start evaluating the model, a process referred to by Salama & El-Gohary (2016) as feature size selection. Meaning, how many independent variables can be used as a starting point for the model to avoid a computational error in the regression model (Daoud, 2017) by creating a smaller amount of predominant feature variables (Hua et al., 2007). This was done by determining the number of highly correlated variables in the Pearson correlation matrix to avoid having too much multicollinearity that prohibits the model from running.

3.3 Stage 3: Feature Selection

Stage 2 resulted in the number of variables for conducting MLR; Stage 3 was conducted to select from those variables. This was done through feature selection, which is simply a reduction of the independent variables to ones that would yield a high-performing model without overfitting (Ashtab & Ryoo, 2022). Feature selection methods can either be classified as "wrapper methods" or "filter methods". Wrapper methods provide the best-performing subset of variables through machine learning, while filter methods rank the variables using statistical tests conducted with raw information (Idowu & Lam, 2020). This research has utilized both of the aforementioned methods since there are varying opinions regarding the superiority of either method. Nnamoko et al. (2014) stated that wrapper methods are superior due to their machine-learning abilities, while Hastie et al. (2009) and Bolón-Canedo et al. (2013) have criticized wrapper approaches due to learning bias and generalization ability.

The filter method chosen was the chi-squared technique due to its promising effects on the performance of subsequent models (Jootoo & Lattanzi, 2017) and due to its application in construction research (Bidgoli & Naseriparsa, 2012; Salama & El-Gohary, 2016). The chi-squared technique involves ranking the independent variables by their chi-squared statistics ($\chi 2$), which is a measure that indicates how much the observed counts of a particular independent variable deviate from the expected counts if the independent and dependent variables are not related (Jootoo & Lattanzi, 2017). Further, the wrapper method chosen is the recursive feature elimination (RFE) method due to its high computational power and its adaptable implementation (Dhal & Azad, 2022) as well as its successful application in a wide variety of construction research (Awada et al., 2021; Chang et al., 2023; Jeong et al., 2024). RFE involves building a machine learning model using all variables and then iteratively removing the least important ones until a stoppage criteria (Thakkar & Lohiya, 2021). The machine learning model used under RFE is a gradient boost classifier (GBC), a technique where iterations rely on adding decision trees and stops when no improvement to the model is reached This technique is known to have a highly reliable performance opposed to other methods (Wong et al., 2021). Overall, by the end of this stage, both methods yielded a ranked list of top-performing variables. Then, the number of variables taken from each was the feature size identified in Stage 2. This has produced two different combinations of factors that would theoretically be of high influence on the rate of safety incidence in the transportation sector.

3.4 Stage 4: Iterative MLR through a Funneling Effect

Two MLR models were created where each model's selected variables/features were the results of the chi-squared and RFE method; further, the number of variables or feature size was limited to that determined in Stage 2. Figure 3.3 provides a flowchart of the process undertaken to create the two MLR models. Initially, the MLR model is run using all variables selected from stages 2 and 3. VIF was used in this research to reduce high multicollinearity. It quantified how much the variance of an estimated regression coefficient was increased due to multicollinearity; high VIF values indicated that a predictor variable may be highly correlated with other predictors in the model (Song & Kroll, 2012). If variables attained VIF values higher than 10, the one with the highest VIF value was eliminated (Idowu & Lam, 2020; Seo et al., 2024). After the variable with the highest VIF was eliminated, the model was run again and the same VIF process was repeated until a model was retrieved where none of the variables had a VIF value greater than 10.

Finally, the model fit was then evaluated using the ANOVA test to make sure that the model significantly explained the outcome variable (Arora et al., 2023). If the test's P-value was less than 0.05, it indicated a good-fit model (Lokesha et al., 2023). The adjusted R-squared and final model variables wer erecorded and models that attained an adjusted R-squared higher than 0.7 were kept (Lokesha et al., 2023). However, if the model did not pass the ANOVA test, all variables were added back and stepwise backward elimination was conducted. Meaning, that the variables were removed according to their effect on the ANOVA P-value (Jelodar et al., 2022) and VIF was evaluated for the variables. This iterative process was applied to the initial two sets of variables, yielding two refined models characterized by distinct variable combinations. Consequently, the authors obtained two models demonstrating strong fits, each representing a unique set of factors significantly influencing safety within the transportation industry. However, eliminating variables in the process meant that some variables were not studied. Since this research requires studying all possible combinations, the authors introduced the funneling effect.

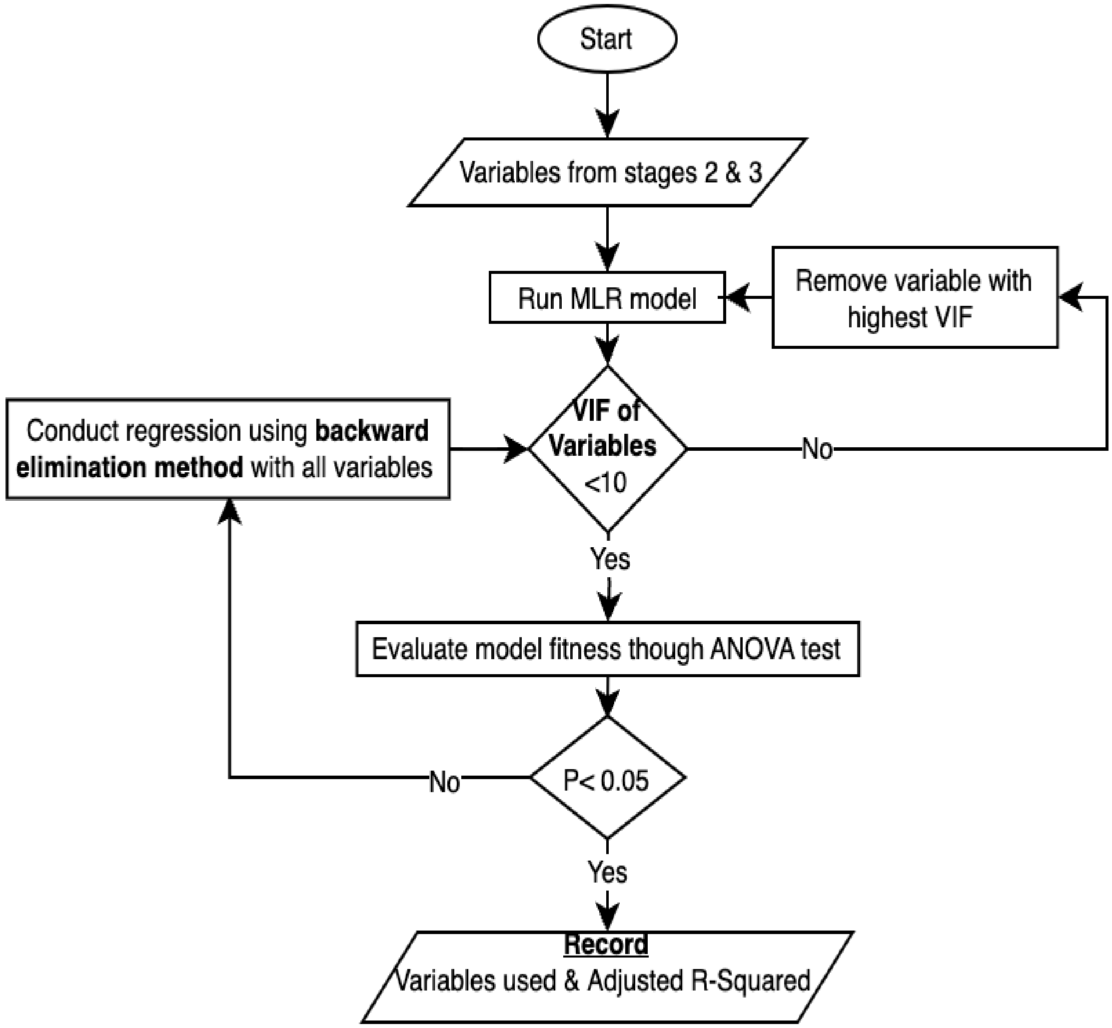

Figure 3.3 MLR Model Process

The funneling effect took all eliminated variables from the first two MLR models and created further good-fit models. The models may not perform better than the first two, but they would satisfy good fit model criteria and thus variable combinations for such models could not be ignored. In fact, the criticality of the upcoming combinations on safety issues may not be as severe but are affecting variables that cannot be ignored. Figure 3.4 provides a representation of the funneling effect using an example of eight independent variables under three different divisions. As seen, the funneling effect started from the results of the first two MLR models created: funnel level 1. From there, unused variables from each division were identified and a list of variable combinations was created. Combinations were created with one variable from every group. This approach mimics being able to study every internal/external factor causing safety issues together. An MLR model was run for every combination using the approach discussed in figure 3.4. The next level of the funnel was reached by repeating the process of removing used variables, creating variable combinations, and conducting MLR until all variables were observed. By the end of this stage, every funnel level would have MLR models with a combination of internal and external variables that affected the rate of transportation safety incidents.

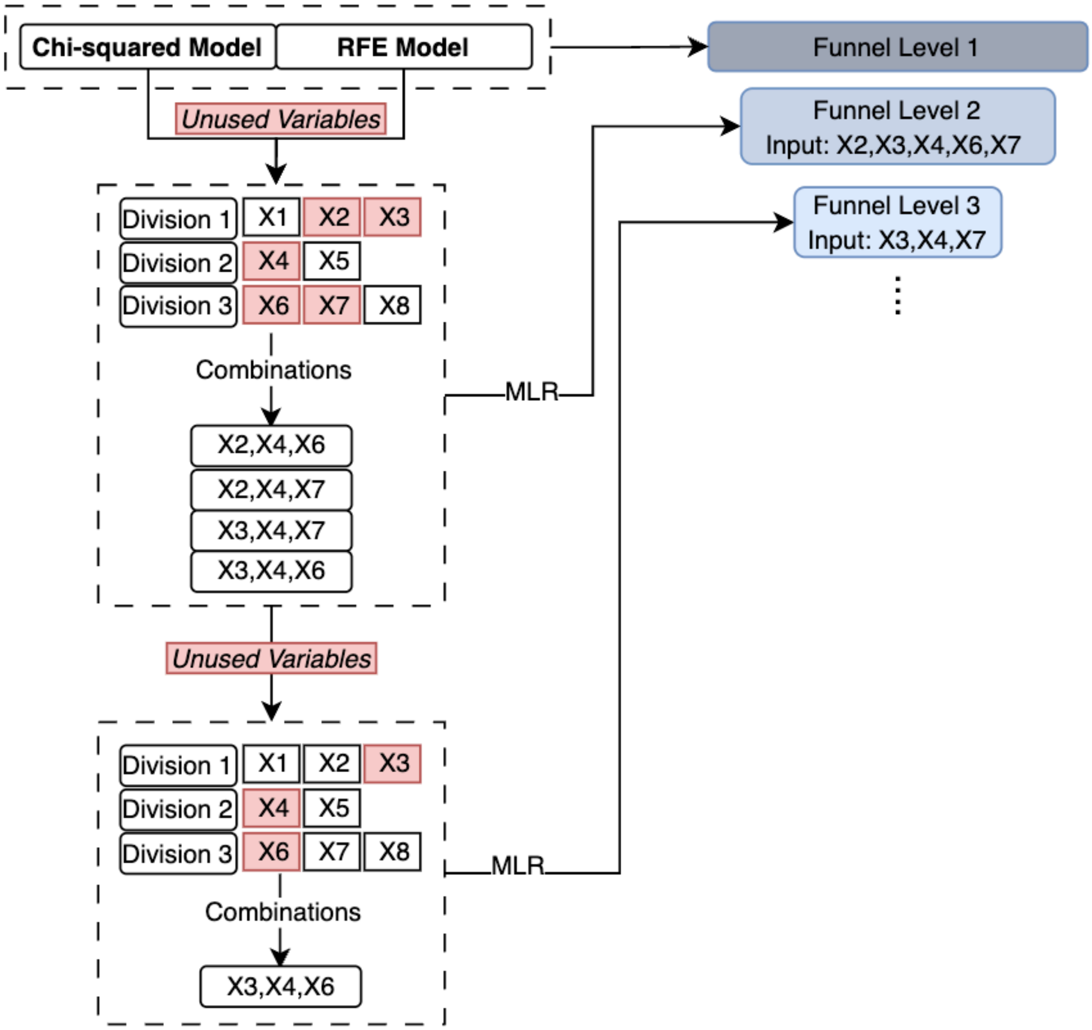

Figure 3.4 Funneling Effect

Chapter 4 Results

4.1 Stage 1: MLR Variables

As discussed in the previous chapter, the total number of divisions was nine. However, before grouping, there were a total of 65 subdivisions. Having 65 independent variables would not be practical; therefore, by adopting the median method, the authors were capable of attaining two groups of subdivisions under every division. The only exception was the division related to the "Event or exposure" and "Occupation". After evaluating the groups created, the authors found that grouping the different events, exposures, and occupations would be too generalized and it is an aspect that must be evaluated with details. Table 4.1 presents the resulting independent and dependent variables. There were 22 independent variables (subdivisions) that belonged to nine different divisions (internal and external factors). They were further used in the MLR stage to observe how influential a group of different factors is on the safety incident rate in the transportation sector.

Table 4.1 MLR Variables

Variable	Division	Subdivision	Notation
Independent Variables	Gender	Men	X1
		Women	X2
	Age	≤44	X3
		>44	X4
	Occupation	Construction and Extraction	X5
		Installation, maintenance, and repair	X6
		Transportation and material moving	X7
	Length of service with employer	<1 year	X8
		≥1 year	X9
	Race or ethnic origin	White	X10
		African American or Hispanic	X11
	Event or exposure	Transportation incidents	X12
		Fires/explosions & Exposure to harmful substances or environments	X13
		Falls, slips, trips	X14
		Contact with object, equipment	X15
		Overexertion and bodily reaction	X16
	Day of Week	Weekend	X17
		Weekday	X18
	Time of Day	12:01 PM - 12:00 AM	X19
		12:01 AM - 12:00 PM	X20
	Hours Worked	≤8	X21
		>8	X22
Dependent Variable	Incident Rate		Y

4.2 Stage 2: Feature Size

This stage involved examining the correlation of the 21 independent variables and yielded the feature size needed to start the MLR model. Figure 4.1 is the resulting correlation matrix where an overall high correlation is present between the variables. The most notable variables are X1 (males in the transportation sector), X3 (workers under the age of 44), X5 (occupation of construction and extraction), X7 (less than one year of service with the employer), X9 (white ethnicity), X15 (events of overexertion and bodily reaction), X17 (working weekdays), X19 (working from 12:00 AM - 12:00 PM), and X20 (working less than eight hours). The combination of variables together did not mean anything at this point regarding their combined impact on safety incidents in the transportation sector. Because their occurrence together in an MLR model did not yield reliable results, they were probably eliminated when observing their VIF. However, standing out as variables with high correlations indicated that their occurrence was not independent of each other. Therefore, while an MLR model was created with the above combination, it could not be ignored as a group of factors that would coexist when a safety accident occurs. This was a limitation of utilizing the MLR to retrieve various combinations, but would be further addressed when all possible combinations were retrieved from the MLR results. For the time being, the feature size was retrieved by counting the number of variables with moderate to negligible correlation with more than one variable, which is any value less than 0.7 (Acharya et al., 2024). The result was a maximum feature size of 11 different variables.

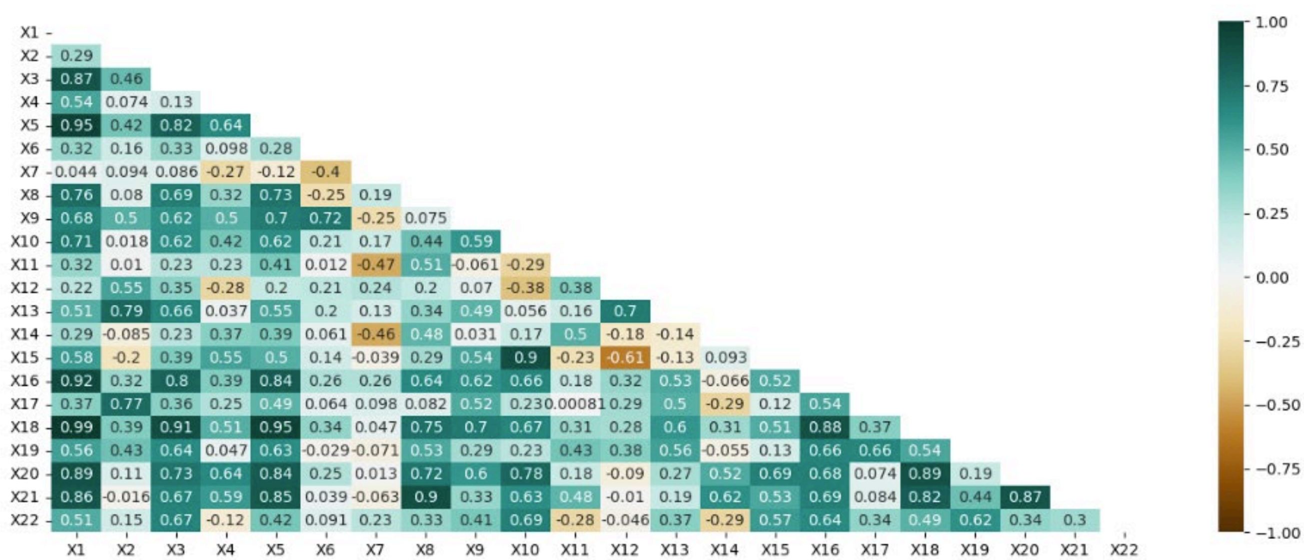

Figure 4.1 Pearson Correlation Matrix

4.3 Stage 3: Feature Elimination

The feature elimination was conducted using the sklearn library in Python for the chi-squared and RFE methods. Table 4.2 shows the results of the ranked variables using both methods. As per the results of stage 2, only the top 11 independent variables from each ranking were utilized to create the first two MLR models. Ranking in both models was different for the seven variables. This was expected since RFE utilizes machine learning to eliminate weak variables while chi-squared utilizes statistical tests by observing every independent variable with the dependent. This means that the chi-squared method may not have accounted for inflation due to high multicollinearity and would require more elimination of variables when creating the MLR model as opposed to the RFE model.

Table 4.2 Feature Selection

Ranking	RFE	Chi-Squared
1	X9	X22
2	X2	X10
3	X10	X9
4	X15	X15
5	X13	X16
6	X17	X12
7	X20	X20
8	X12	X5
9	X6	X1
10	X4	X3
11	X21	X18
12	X14	X2
13	X22	X21
14	X3	X6
15	X5	X14
16	X11	X7
17	X7	X4
18	X8	X13
19	X16	X17
20	X18	X8
21	X19	X19
22	X1	X11

4.4 Stage 4: MLR Results

The final section of the results demonstrates all MLR models attained. Figure 4.2 demonstrates the independent variables used at every funnel level (every rectangle represents a variable in a single division), the MLR models retrieved, and the number of variables within each model: there was a total of three funnel levels. The first level involved 22 independent variables, 11 for each model, and after following the MLR procedure set forth above, two high-performing MLR models were yielded. The eliminated variables were then taken, and an assortment was formulated that encompassed different combinations where one variable was selected from every division as demonstrated in figure 4.2. Twelve different combinations were retrieved and MLR was conducted on all of them. This resulted in three different top-performing MLR models. Finally, the unused variables were used as input into funnel level three, which attained one combination yielding one high-performing MLR model. It is noteworthy that studying one variable from every division provided this methodology with the logical inference through statistical methods that attributes affect the rate of safety incidents in the transportation sector. Meaning, in this context, saying that both females and males together (X1 & X2) with a certain occupation affect the rate of safety incidents would not have much bearing on the required corrective action.

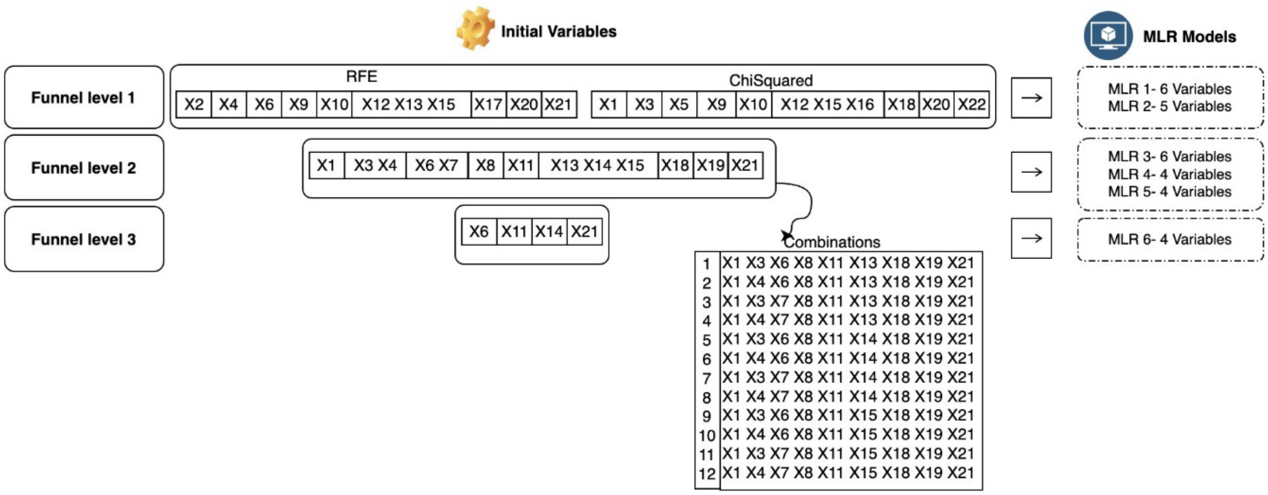

Figure 4.2 Funnel MLR Results

As discussed in the methodology section, the method must be statistically sound by observing the criteria threshold that formulates a good-performing model. Table 4.3 presents the adjusted R-squared, the ANOVA test P-value, and the variables' VIF value for all models. All R values are above 0.7, all P values are below 0.05, all VIF Values are below 10, and all variables have been studied.

Table 4.3 MLR Model Performance

Model	Adjusted R-Squared	ANOVA P-Value	VIF Values	
MLR 1	0.826	0.038	X2	4.08
			X9	2.654
			X10	5.379
			X12	2.104
			X17	3.67
			X20	4.189
MLR 2	0.963	0.001	X5	4.843
			X9	2.115
			X10	2.704
			X16	5.112
			X22	2.565
MLR 3	0.965	0.005	X1	5.415
			X3	6.1
			X7	1.169
			X8	2.818
			X13	2.113
			X19	2.011
MLR 4	0.83	0.009	X7	1.064
			X8	2.46
			X15	1.389
			X18	2.921
MLR 5	0.778	0.017	X1	4.111
			X4	1.639
			X8	2.471
			X13	1.529
MLR 6	0.745	0.024	X6	1.004
			X11	1.428
			X14	1.788
			X21	1.744

Finally, to put the variables into a divisional context, the various combinations are demonstrated in figure 4.3 from most to least impacting on the safety incidents based on the R-value for the MLR model. The upcoming chapter encompasses a detailed discussion relevant to the individual combinations.

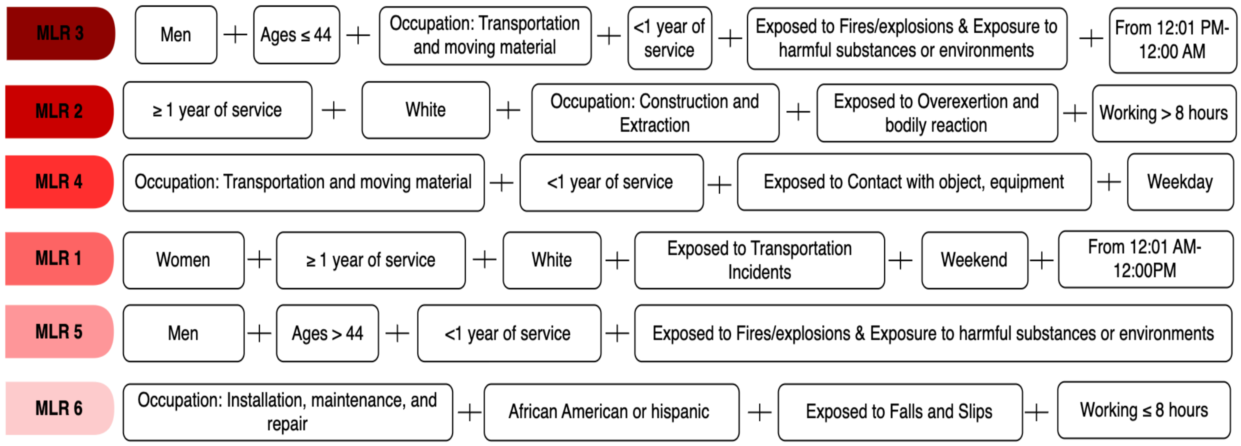

Figure 4.3 Factors Impacting Safety Rate

Chapter 5 Deciphering Safety Challenges

This section is divided into separate discussions of every MLR from most to least impacting on safety incidents based on the R-value of the model.

5.1 MLR 3

The first combination involves individual factors recognized as highly influential as mentioned separately in the literature but their coexistence is yet to be deemed as critical. Namian et al. (2022) have recently concluded that it is not enough to study age as the sole variable impacting safety, meaning one cannot make the judgment of a worker being more prone to safety incidents based on their inexperience alone. This study highlights that men under the age of 44 employed in transportation and material moving roles face a heightened risk of fire and explosions, particularly those with less than one year of service with their employer. The combination of relatively inexperienced workers due to age and due to a short length of time working with an employer can be very critical (Choi et al., 2020).

A person's working hours can have an effect on safety, which largely depends on their occupation. The BLS (2024) provides an outlook on material moving machine operators' work schedules where "materials are shipped around the clock, some work overnight shifts." This can lead to fatigue-induced errors and amplify the risk of accidents (Koc et al., 2023). Consequently, morning shifts fall in the time frame identified as most likely for fire accidents to occur on construction sites (J.S. Kim & Kim, 2018). The nature of the work, often involving handling flammable materials or operating machinery prone to sparking, also increases the likelihood of such incidents. The nature of work in transportation and material moving roles significantly contributes to the heightened risk of fire and explosions in the workplace. These occupations frequently involve handling flammable materials and operating machinery prone to sparking, thus increasing the likelihood of such incidents (Hassanain et al., 2022). Workers in these roles may regularly come into contact with substances such as gasoline, propane, or industrial chemicals, which pose inherent fire hazards. Additionally, the operation of heavy machinery, such as forklifts or conveyor systems, can generate sparks or friction that are a source of heat and may ignite flammable materials in the vicinity (Campbell, 2023). The combination of these factors creates a volatile working environment where the potential for fire and explosions is heightened, necessitating stringent safety protocols and risk management strategies to mitigate these risks effectively. The combination of youthful exuberance, limited tenure with the employer, and a demanding schedule underscores the importance of robust safety training, vigilant supervision, and comprehensive risk assessment protocols to safeguard these workers against fire and explosion hazards while operating equipment and moving material.

5.2 MLR 2

The second most critical combination focuses on the portion of the workforce that has had more experience working for a certain employer. This may add more familiarity to the work and procedures which can reduce the risk of exposure to safety incidents. However, as individuals accumulate experience and familiarity with their job responsibilities, workplace environment, and colleagues, they often develop a greater sense of competence and comfort in their roles. This increased confidence can translate into a willingness to undertake extended work hours or more demanding shifts, which could eventually lead to overexertion of the worker (Caruso et al., 2006).

Workers in the transportation sector, particularly those with more than one year of service with an employer, face significant risks of overexertion and bodily reaction, especially when required to work extended shifts exceeding eight hours. Research indicates that prolonged hours of physical labor, common in transportation roles such as truck driving or loading and unloading freight, significantly increase the likelihood of musculoskeletal injuries due to overexertion (Everett, 1999). Moreover, the demanding nature of the job, which often involves repetitive tasks and heavy lifting, exacerbates the risk of strains, sprains, and other bodily reactions among workers (Dong et al., 2019; Smith et al., 2023; X. Wang et al., 2017). The aforementioned are found under the feature of construction and extraction occupation in this model which are occupations that lead to overexertion (D. Wang et al., 2015).

The above-mentioned hazards are particularly pronounced for white workers in the transportation sector (Dong et al., 2019). This finding is stimulating since among laborers, white workers are more likely to attain a higher level of education (BLS, 2021) and are most likely to exert less physical activity at work (Saffer et al., 2013). However, studies have highlighted disparities in occupational safety outcomes based on race, with white workers often showing more emotional disparity in injury-related situations (Bhandari & Hallowell, 2017), more susceptible to physical fatigue due to heat (Karthick et al., 2022), and are more likely to feel the need to rigorously prove themselves among other ethnicities, which increases their likelihood to work more hours and lead to overexertion (Paap, 2006). Addressing these disparities requires a comprehensive approach that considers the specific needs and experiences of white workers, including implementing ergonomic interventions, providing adequate rest breaks during extended shifts, and promoting a culture of safety and well-being in the workplace. While the mention of working more than eight hours and overexertion may be anticipated, the newly added perspective within this research is its criticality of occurrence when workers are of a white race and have worked with the employer for more than a year.

5.3 MLR 4

This model contains four different features, two of which are the same as the first model: less than one year of experience and having an occupation involving transportation and moving material. While these features are still a critical combination with MLR 3, they become less critical when coupled with working weekdays and workers exposed to contact with objects and equipment. Nevertheless, it is important to shed light on the third most critical combination of features. The fast-paced nature of weekday operations in transportation and movement roles may lead to rushed or hurried tasks, further heightening the likelihood of workplace incidents (Han et al., 2014). These individuals often find themselves in roles that involve direct contact with heavy machinery, such as forklifts, cranes, or conveyor systems (Cho & Gai, 2014). The lack of experience with the employer coupled with the demands of the job increases the risk of accidents and injuries related to equipment malfunctions, mishandling of materials, or improper use of machinery. Given the lack of familiarity with safety protocols and proper handling procedures among workers with less than one year of service, there is an urgent need for comprehensive training programs and stringent safety measures to mitigate the risks.

5.4 MLR 1

This model presents the female gender as opposed to males used in other models. Females compose only 10.9% of the entire construction sector workforce (BLS, 2022), yet gender is still an important feature of the fourth most important combination of variables. This disproportion in the number of females versus the serious effect on safety incidents is another noteworthy aspect of this research's results. Females are more prone to transportation incidents(Tork, 2008). Hasan & Kamardeen (2022) attribute this to a "higher number of females on current construction sites working as road flaggers."

Moreover, these incidents are exacerbated from 12:00 am to 12:00 pm on a weekend since drivers do not expect work zones to be occupied at night time, causing them to speed and have less visibility (K. Zhang & Hassan, 2019). Weekend crashes were also found to be more severe (Osman et al., 2019), which can be attributed to an increase in drunk drivers (W. Zhang & Zhang, 2020). Additionally, the disruption of circadian rhythms due to working overnight may exacerbate these risks, as the body's natural tendency to rest conflicts with the demands of the job (Barger et al., 2012).

Finally, females working during the weekend after midnight who are being exposed to transportation incidents contain two previously discussed features of having more than one year of service with an employer and being white. The repetition of this combination emphasizes that white workers working with an employer for an extended period in the construction industry can potentially diminish a worker's cautiousness. Long-tenured employees may become overly familiar with their tasks and surroundings, leading to complacency and a reduced tendency to adhere strictly to safety protocols. This phenomenon, known as "risk homeostasis", suggests that individuals adjust their behavior based on perceived levels of risk, potentially taking greater risks when they feel overly confident or accustomed to their work environment (McKinnon, 2012). Prolonged exposure to the same tasks and conditions may lead to a false sense of security, causing workers to underestimate potential dangers or overlook warning signs, further compromising their cautiousness.

5.5 MLR 5

The relationship between injury incidents and age is inversely proportional where younger workers are more prone to incidents (Chen et al., 2016). This lies in consensus with the findings of this model where ages greater than 44 are presented as less critical to safety in the transportation sector as opposed to the younger age group that was present in the most critical model. As mentioned, accounting for age alone is unreliable, so this model attains three other variables discussed under the first model: males with less than one year of service who are exposed to fires or explosions and harmful substances. The repetition reiterates how critical this combination is but was ranked less critical due to the age group feature. These individuals, often new to the job or recently hired, may lack the necessary training and familiarity with safety protocols. The hierarchical structure often observed in construction crews, where experienced workers of an older age group in the company may hold positions of authority over newer recruits, can create barriers to effective communication and hinder the transfer of essential safety knowledge and skills.

5.6 MLR 6

Workers in the transportation sector, particularly those engaged in installation, maintenance, and repair tasks, who are of Hispanic or Black ethnicity and have worked a shift of fewer than eight hours, are disproportionately exposed to risks of falls and slips. The first notable aspect would be the nature of the event occurring when workers have worked less than an eight-hour shift. Chan et al. (2008) examined workers engaged in repair, maintenance, and alterations additions while working from heights. One of the conclusions was that time of day was an important factor, and falls seemed to occur more frequently during the afternoon from 14:01-16:00. This means, that the worker would not have had a chance to complete an eight-hour shift (Kines, 2002). While this type of exposure to an event may be serious within the construction industry (Hu et al., 2011), it may not be abundant in transportation (Lipscomb et al., 2006).

The last feature present within this combination would be workers of African American or Hispanic ethnicity. This was because falls were found to be caused by negligence which is among the top four reasons for Hispanic workers' causes of safety incidents (Loayza Chahuayo, 2011). While research about Hispanics in construction is abundant, the African American ethnicity does not seem to play a major role in construction literature. In fact, Brigham et al. (2012) have highlighted the lack of the African American community working in the construction sector. This does not negate that the presence of this ethnicity has contributed to causing fall and slip safety incidents in the transportation sector. While it is not clear why, it must be prevented when included with the amalgam of factors discussed in this model.

Chapter 6 Conclusions and Recommendations

In conclusion, this research highlights the critical combinations of variables influencing safety incidents in the transportation sector, stressing the need for tailored interventions to effectively mitigate risks. Utilizing a methodological approach involving data retrieval from publicly available sources, factor elimination, and funnel-based MLR analysis, five significant combinations of factors impacting safety incident rates were identified and discussed. The practical contribution of this research lies in its identification and analysis of critical combinations of variables that significantly impact safety incidents in the transportation sector, thereby informing targeted interventions and preventive measures to enhance workplace safety. State DOTs would benefit from evading such critical combinations to observe lower safety incident rates. While considering all factors forming one combination is important, there were factors identified to have a higher impact:

- Newly hired individuals often face challenges due to a lack of training, particularly when they have less than one year of experience with the employer. This limited tenure may result in insufficient time for comprehensive onboarding and safety training programs, leaving workers ill-prepared to navigate the hazards of their new roles.
- Older individuals entering a new job may encounter safety issues stemming from hierarchical clashes with younger colleagues. This intergenerational dynamic can create tensions and communication barriers, particularly in workplaces where younger workers may hold supervisory roles or possess advanced technical skills.
- Women in the sector, particularly those placed as flagmen, may face unique safety challenges due to the nature of their tasks. Being positioned in potentially hazardous

environments exposes them to risks associated with passing vehicles and heavy machinery.

- White men with over one year of service with an employer may face challenges related to complacency and overconfidence. Prolonged tenure can lead to a false sense of security, causing workers to underestimate dangers or overlook safety protocols.
- The time of day for Hispanic workers plays a significant role in falls within the transportation sector, necessitating heightened safety measures such as increased supervision during the afternoon to mitigate risks.
- The underrepresentation of African Americans in safety research within the transportation sector is a critical aspect to consider. Limited research fails to capture the full extent of their experiences and the factors contributing to safety incidents, potentially leading to disparities in safety outcomes and ineffective interventions.

The above underscores the importance of addressing both internal and external dynamics to promote a culture of safety excellence. The research contributes substantially to transportation safety knowledge by introducing a novel methodology and offering actionable insights for industry stakeholders, policymakers, and safety professionals. By deciphering the interplay of internal and external factors contributing to occupational injuries, the research provides practical insights to enhance worker well-being and safety culture. Additionally, identifying critical combinations such as challenges faced by new hires, hierarchical clashes, and specific safety risks for women and minority groups emphasizes the need for diverse demographic considerations in safety interventions. Moving forward, State DOTs must recognize these findings' significance and implement targeted interventions to mitigate safety risks, prioritizing worker well-being and adopting multifaceted safety management approaches for a safer transportation work environment nationwide. In this endeavor, practical contributions to State DOTs are paramount. State DOTs can facilitate partnerships between government agencies, industry stakeholders, and advocacy groups to address the extracted safety concerns collaboratively ensuring the well-being of all workers in the industry.

References

Acharya, S., Khandegar, V., & Sharma, S. K. (2024). Groundwater Quality Assessment of Newly Developed Dwarka Region, New Delhi, India. Journal of Hazardous, Toxic, and Radioactive Waste, 28(2), 04023046. https://doi.org/10.1061/JHTRBP.HZENG-1265

Al-Bayati, A. J., Ali, M., & Nnaji, C. (2023). Managing Work Zone Safety during Road Maintenance and Construction Activities: Challenges and Opportunities. Practice Periodical on Structural Design and Construction, 28(1), 04022068. https://doi.org/10.1061/PPSCFX.SCENG-1212

Al-Bayati, A. J., Eiris, R., Hernandez, G. O., & Al-Bayati, M. A. (2023). Addressing Social Inequity in Construction Safety Research: Personal Protective Equipment Focus. Journal of Construction Engineering and Management, 149(12), 04023127. https://doi.org/10.1061/JCEMD4.COENG-13864

Alexander, D., Hallowell, M., & Gambatese, J. (2017). Precursors of Construction Fatalities. II: Predictive Modeling and Empirical Validation. *Journal of Construction Engineering and Management*, *143*(7), 04017024. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001297

Alqatawna, A., Rivas Álvarez, A. M., & García-Moreno, S. S.-C. (2021). Comparison of Multivariate Regression Models and Artificial Neural Networks for Prediction Highway Traffic Accidents in Spain: A Case Study. *Transportation Research Procedia*, *58*, 277–284. https://doi.org/10.1016/j.trpro.2021.11.038

Al-Shabbani, Z., Sturgill, R., & Dadi, G. (2018). *Towards Improving Safety Performance of Transportation Maintenance Workers through a Pre-Task Safety Talk Program*. 624–634. https://doi.org/10.1061/9780784481288.061

Ammar, A., & Dadi, G. (2023). *Roadmap for Redesigning Specialized Safety Training and Tracking for Highway Construction and Maintenance Personnel*. 683–694. https://doi.org/10.1061/9780784484883.059

Arditi, D., Ayrancioglu, M. A., & Shi, J. (2004). Effectiveness of Safety Vests in Nighttime Highway Construction. *Journal of Transportation Engineering*, *130*(6), 725–732. https://doi.org/10.1061/(ASCE)0733-947X(2004)130:6(725)

Arora, J., Kar, S., & Jha, K. N. (2023). Assessment of the Significance of Identified Attributes Affecting the Rebar-Fixing Productivity Using Multiple Regression. *Journal of Construction Engineering and Management*, *149*(6), 04023039. https://doi.org/10.1061/JCEMD4.COENG-12861

Ashtab, M., & Ryoo, B. Y. (2022). Predicting Construction Workforce Demand Using a Combination of Feature Selection and Multivariate Deep-Learning Seq2seq Models. *Journal of Construction Engineering and Management*, *148*(12), 04022136. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002414

Assaad, R., & El-adaway, I. H. (2021a). Determining Critical Combinations of Safety Fatality Causes Using Spectral Clustering and Computational Data Mining Algorithms. *Journal of Construction Engineering and Management*, *147*(5), 04021035. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002040

Assaad, R., & El-adaway, I. H. (2021b). Determining Critical Combinations of Safety Fatality Causes Using Spectral Clustering and Computational Data Mining Algorithms. *Journal of Construction Engineering and Management*, *147*(5), 04021035. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002040

Awada, M., Srour, F. J., & Srour, I. M. (2021). Data-Driven Machine Learning Approach to Integrate Field Submittals in Project Scheduling. *Journal of Management in Engineering*, *37*(1), 04020104. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000873

Ayhan, B. U., & Tokdemir, O. B. (2020a). Accident Analysis for Construction Safety Using Latent Class Clustering and Artificial Neural Networks. *Journal of Construction Engineering and Management*, *146*(3), 04019114. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001762

Ayhan, B. U., & Tokdemir, O. B. (2020b). Accident Analysis for Construction Safety Using Latent Class Clustering and Artificial Neural Networks. *Journal of Construction Engineering and Management*, *146*(3), 04019114. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001762

Bader, S. M., Abotaleb, I., & Hosny, O. (2024). *A VR-Based Safety Training Program for General and Enabling Works in Construction Projects*. 70–78. https://doi.org/10.1061/9780784485231.009

Balali, F., Nouri, J., Nasiri, A., & Zhao, T. (2020). Data Analytics. In F. Balali, J. Nouri, A. Nasiri, & T. Zhao (Eds.), *Data Intensive Industrial Asset Management: IoT-based Algorithms and Implementation* (pp. 105–113). Springer International Publishing. https://doi.org/10.1007/978-3-030-35930-0_7

Barger, L. K., Ogeil, R. P., Drake, C. L., O'Brien, C. S., Ng, K. T., & Rajaratnam, S. M. W. (2012). Validation of a Questionnaire to Screen for Shift Work Disorder. *Sleep*, *35*(12), 1693–1703. https://doi.org/10.5665/sleep.2246

Benesty, J., Chen, J., & Huang, Y. (2008). On the Importance of the Pearson Correlation Coefficient in Noise Reduction. *IEEE Transactions on Audio, Speech, and Language Processing*, *16*(4), 757–765. https://doi.org/10.1109/TASL.2008.919072

Bhandari, S., & Hallowell, M. R. (2017). Emotional Engagement in Safety Training: Impact of Naturalistic Injury Simulations on the Emotional State of Construction Workers. *Journal of Construction Engineering and Management*, *143*(12), 04017090. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001405

Bidgoli, A. M., & Naseriparsa, M. (2012). A Hybrid Feature Selection by Resampling, Chi-squared abd Consistency Evaluation Techniques. *World Academy of Science, Engineering and Technology*, 276–285.

BLS. (2021). *Labor force characteristics by race and ethnicity, 2020*. https://www.bls.gov/opub/reports/race-and-ethnicity/2020/home.htm

BLS. (2022). *Women in the labor force: A databook*. https://www.bls.gov/opub/reports/womens-databook/2021/home.htm

BLS. (2023a). *Incidence rates of nonfatal occupational injuries and illnesses by industry and case types, 2022*. https://www.bls.gov/web/osh/table-1-industry-rates-national.htm

BLS. (2023b). *Industries at a Glance: Construction: NAICS 23 : U.S. Bureau of Labor Statistics*. https://www.bls.gov/iag/tgs/iag23.htm

BLS. (2023c). *Survey of Occupational Injuries and Illnesses Data: U.S. Bureau of Labor Statistics*. https://www.bls.gov/iif/overview/soii-overview.htm

BLS. (2024). *Material Moving Machine Operators: Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics*. https://www.bls.gov/ooh/transportation-and-material-moving/material-moving-machine-operators.htm#tab-2

Bobadilla, L., Mostafavi, A., Carmenate, T., & Bista, S. (2014). *Predictive Assessment and Proactive Monitoring of Struck-By Safety Hazards in Construction Sites: An Information Space Approach*. 989–996. https://doi.org/10.1061/9780784413616.123

Bolón-Canedo, V., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2013). A review of feature selection methods on synthetic data. *Knowledge and Information Systems*, *34*(3), 483–519. https://doi.org/10.1007/s10115-012-0487-8

Bortey, L., Edwards, D. J., Roberts, C., & Rillie, I. (2022). A Review of Safety Risk Theories and Models and the Development of a Digital Highway Construction Safety Risk Model. *Digital*, *2*(2), Article 2. https://doi.org/10.3390/digital2020013

Brigham, D., Nobe, M., & Glick, S. (2012). *Increasing African American Participation in the Construction Industry*. 48th ASC Annual International Conference Proceedings.

Campbell, R. (2023, October 1). *Fires in Structures Under Construction*. National Fire Protection Association. https://www.nfpa.org/education-and-research/research/nfpa-research/fire-statistical-reports/fires-in-structures-under-construction

Caruso, C. C., Bushnell, T., Eggerth, D., Heitmann, A., Kojola, B., Newman, K., Rosa, R. R., Sauter, S. L., & Vila, B. (2006). Long working hours, safety, and health: Toward a national research agenda. *American Journal of Industrial Medicine*, *49*(11), 930–942. https://doi.org/10.1002/ajim.20373

Chammout, B., Ahmed, M. O., El-adaway, I., & Lieser, W. (2024). A Holistic Approach to Exploring the Root Factors of Work Zone Accidents. *Journal of Management in Engineering*, *40*(1), 04023062. https://doi.org/10.1061/JMENEA.MEENG-5729

Chan, A. P. C., Wong, F. K. W., Chan, D. W. M., Yam, M. C. H., Kwok, A. W. K., Lam, E. W. M., & Cheung, E. (2008). Work at Height Fatalities in the Repair, Maintenance, Alteration, and Addition Works. *Journal of Construction Engineering and Management*, *134*(7), 527–535. https://doi.org/10.1061/(ASCE)0733-9364(2008)134:7(527)

Chang, T., Lee, G., & Chi, S. (2023). Development of an Optimized Condition Estimation Model for Bridge Components Using Data-Driven Approaches. *Journal of Performance of Constructed Facilities*, *37*(3), 04023013. https://doi.org/10.1061/JPCFEV.CFENG-4359

Chen, Y., McCabe, B., & Hyatt, D. (2016). *Safety and Age: A Longitudinal Study of Ontario Construction Workers*. 2957–2966. https://doi.org/10.1061/9780784479827.294

Chiang, Y.-H., Wong, F. K.-W., & Liang, S. (2018). Fatal Construction Accidents in Hong Kong. *Journal of Construction Engineering and Management*, *144*(3), 04017121. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001433

Cho, Y. K., & Gai, M. (2014). Projection-Recognition-Projection Method for Automatic Object Recognition and Registration for Dynamic Heavy Equipment Operations. *Journal of Computing in Civil Engineering*, *28*(5), A4014002. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000332

Choe, S., & Leite, F. (2017). Assessing Safety Risk among Different Construction Trades: Quantitative Approach. *Journal of Construction Engineering and Management*, *143*(5), 04016133. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001237

Choi, J., Gu, B., Chin, S., & Lee, J.-S. (2020). Machine learning predictive model based on national data for fatal accidents of construction workers. *Automation in Construction*, *110*, 102974. https://doi.org/10.1016/j.autcon.2019.102974

Cooksey, R. W. (2020). Descriptive Statistics for Summarising Data. In R. W. Cooksey (Ed.), *Illustrating Statistical Procedures: Finding Meaning in Quantitative Data* (pp. 61–139). Springer. https://doi.org/10.1007/978-981-15-2537-7_5

Cote, C. (2021, October 19). 4 Types of Data Analytics to Improve Decision-Making. *Harvard Business School Online's Business Insights BlogBusiness Insights Blog*. https://online.hbs.edu/blog/post/types-of-data-analysis

Dang, H., Serne, J., & Tafazzoli, M. (2024). *Virtual Reality Safety Training Assessment in Construction Management and Safety and Health Management Programs*. 28–35. https://doi.org/10.1061/9780784485231.004

Daoud, J. I. (2017). Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series*, *949*(1), 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Das, S., Dutta, A., Tamakloe, R., & Khan, M. N. (2023). Analyzing the time-varying patterns of contributing factors in work zone-related crashes. *Journal of Transportation Safety & Security*, *0*(0), 1–28. https://doi.org/10.1080/19439962.2023.2246020

Dhal, P., & Azad, C. (2022). A comprehensive survey on feature selection in the various fields of machine learning. *Applied Intelligence*, 1–39.

Dong, X. S., Betit, E., Dale, A. M., Barlet, G., & Wei, Q. (2019). *Trends of musculoskeletal disorders and interventions in the construction industry* (CPWR Quarterly Data Report). https://stacks.cdc.gov/view/cdc/86273

Dong, X. S., Choi, S. D., Borchardt, J. G., Wang, X., & Largay, J. A. (2013). Fatal falls from roofs among U.S. construction workers. *Journal of Safety Research*, *44*, 17–24. https://doi.org/10.1016/j.jsr.2012.08.024

Esmaeili, B., & Hallowell, M. (2013). Integration of safety risk data with highway construction schedules. *Construction Management and Economics*, *31*(6), 528–541. https://doi.org/10.1080/01446193.2012.739288

Everett, J. G. (1999). Overexertion Injuries in Construction. *Journal of Construction Engineering and Management*, *125*(2), 109–114. https://doi.org/10.1061/(ASCE)0733-9364(1999)125:2(109)

Fass, S., Yousef, R., Liginlal, D., & Vyas, P. (2017). Understanding causes of fall and struck-by incidents: What differentiates construction safety in the Arabian Gulf region? *Applied Ergonomics*, *58*, 515–526. https://doi.org/10.1016/j.apergo.2016.05.002

Fu, H., Xia, Z., Tan, Y., & Guo, X. (2024). Influence of Cues on the Safety Hazard Recognition of Construction Workers during Safety Training: Evidence from an Eye-Tracking Experiment. *Journal of Civil Engineering Education*, *150*(1), 04023009. https://doi.org/10.1061/JCEECD.EIENG-1882

Gao, Y., Gonzalez, V. A., & Yiu, T. W. (2019). The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Computers & Education*, *138*, 101–115. https://doi.org/10.1016/j.compedu.2019.05.003

Garg, S., & Misra, S. (2021). Understanding the components and magnitude of the cost of quality in building construction. *Engineering, Construction and Architectural Management*, *29*(1), 26–48. https://doi.org/10.1108/ECAM-08-2020-0642

Ghodrati, N., Yiu, T. W., Wilkinson, S., & Shahbazpour, M. (2018). A new approach to predict safety outcomes in the construction industry. *Safety Science*, *109*, 86–94. https://doi.org/10.1016/j.ssci.2018.05.016

Ghosh, S., Nourihamedani, M., Reyes, M., & Snyder, L. (2023). Association Between Leading Indicators of Safety Performance in Construction Projects. *International Journal of*

Construction Education and Research, *0*(0), 1–15. https://doi.org/10.1080/15578771.2023.2195209

Gibb, A., Lingard, H., Behm, M., & Cooke, T. (2014). Construction accident causality: Learning from different countries and differing consequences. *Construction Management and Economics*, *32*(5), 446–459. https://doi.org/10.1080/01446193.2014.907498

Guo, S., Tang, B., Liang, K., Zhou, X., & Li, J. (2021). Comparative Analysis of the Patterns of Unsafe Behaviors in Accidents between Building Construction and Urban Railway Construction. *Journal of Construction Engineering and Management*, *147*(5), 04021027. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002013

Hadikusumo, B. H. W., Jitwasinkul, B., & Memon, A. Q. (2017). Role of Organizational Factors Affecting Worker Safety Behavior: A Bayesian Belief Network Approach. *Procedia Engineering*, *171*, 131–139. https://doi.org/10.1016/j.proeng.2017.01.319

Hallowell, M., Esmaeili, B., & Chinowsky, P. (2011). Safety risk interactions among highway construction work tasks. *Construction Management and Economics*, *29*(4), 417–429. https://doi.org/10.1080/01446193.2011.552512

Han, S., Saba, F., Lee, S., Mohamed, Y., & Peña-Mora, F. (2014). Toward an understanding of the impact of production pressure on safety performance in construction operations. *Accident Analysis & Prevention*, *68*, 106–116. https://doi.org/10.1016/j.aap.2013.10.007

Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer New York. https://doi.org/10.1007/978-1-4757-3462-1

Harris, W., Trueblood, A. B., Brooks, R. D., & Brown, S. (2022). *Fatal and nonfatal transportation injuries in the construction industry, 2011-2020* (cdc:121119). https://stacks.cdc.gov/view/cdc/121119

Hasan, A., & Kamardeen, I. (2022). Occupational Health and Safety Barriers for Gender Diversity in the Australian Construction Industry. *Journal of Construction Engineering and Management*, *148*(9), 04022100. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002352

Hassanain, M. A., Al-Harogi, M., & Ibrahim, A. M. (2022). Fire Safety Risk Assessment of Workplace Facilities: A Case Study. *Frontiers in Built Environment*, *8*. https://www.frontiersin.org/articles/10.3389/fbuil.2022.861662

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer. https://doi.org/10.1007/978-0-387-84858-7

Hu, K., Rahmandad, H., Smith-Jackson, T., & Winchester, W. (2011). Factors influencing the risk of falls in the construction industry: A review of the evidence. *Construction Management and Economics*, *29*(4), 397–416. https://doi.org/10.1080/01446193.2011.558104

Hu, Z., Chan, W. T., Hu, H., & Xu, F. (2023). Cognitive Factors Underlying Unsafe Behaviors of Construction Workers as a Tool in Safety Management: A Review. *Journal of Construction Engineering and Management*, *149*(3), 03123001. https://doi.org/10.1061/JCEMD4.COENG-11820

Hua, X. G., Ni, Y. Q., Ko, J. M., & Wong, K. Y. (2007). Modeling of Temperature–Frequency Correlation Using Combined Principal Component Analysis and Support Vector Regression Technique. *Journal of Computing in Civil Engineering*, *21*(2), 122–135. https://doi.org/10.1061/(ASCE)0887-3801(2007)21:2(122)

Idowu, O. S., & Lam, K. C. (2020). Conceptual Quantities Estimation Using Bootstrapped Support Vector Regression Models. *Journal of Construction Engineering and Management*, *146*(4), 04020018. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001780

Jelodar, M. B., Yiu, T. W., & Wilkinson, S. (2022). Empirical Modeling for Conflict Causes and Contractual Relationships in Construction Projects. *Journal of Construction Engineering and Management*, *148*(5), 04022017. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002263

Jeon, J., & Cai, H. (2022). Multi-class classification of construction hazards via cognitive states assessment using wearable EEG. *Advanced Engineering Informatics*, *53*, 101646. https://doi.org/10.1016/j.aei.2022.101646

Jeon, J., & Cai, H. (2023). Wearable EEG-based construction hazard identification in virtual and real environments: A comparative study. *Safety Science*, *165*, 106213. https://doi.org/10.1016/j.ssci.2023.106213

Jeong, E., Jang, J., Lee, S., Ahn, S., & Kim, T. W. (2024). Quantitative Analysis of the Factors Influencing Field Installation Time for Precast Concrete Building Components: An Empirical Study. *Journal of Management in Engineering*, *40*(3), 04024010. https://doi.org/10.1061/JMENEA.MEENG-5816

Jootoo, A., & Lattanzi, D. (2017). Bridge Type Classification: Supervised Learning on a Modified NBI Data Set. *Journal of Computing in Civil Engineering*, *31*(6), 04017063. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000712

Karthick, S., Kermanshachi, S., & Loganathan, K. (2022). *Occupational Fatigue and Physical Health of Construction Workers in Extreme Hot Weather*. 259–269. https://doi.org/10.1061/9780784484609.028

Karthick, S., Kermanshachi, S., Rouhanizadeh, B., & Namian, M. (2021). *Short- and Long-Term Health Challenges of Transportation Workforce due to Extreme Weather Conditions*. 39–51. https://doi.org/10.1061/9780784483787.005

Kim, H., Seong, J., & Jung, H.-J. (2023). Real-Time Struck-By Hazards Detection System for Small- and Medium-Sized Construction Sites Based on Computer Vision Using Far-Field Surveillance Videos. *Journal of Computing in Civil Engineering*, *37*(6), 04023028. https://doi.org/10.1061/JCCEE5.CPENG-5238

Kim, J.-S., & Kim, B.-S. (2018). Analysis of Fire-Accident Factors Using Big-Data Analysis Method for Construction Areas. *KSCE Journal of Civil Engineering*, *22*(5), 1535–1543. https://doi.org/10.1007/s12205-017-0767-7

Kim, Y. A., Ryoo, B. Y., Kim, Y.-S., & Huh, W. C. (2013). Major Accident Factors for Effective Safety Management of Highway Construction Projects. *Journal of Construction Engineering and Management*, *139*(6), 628–640. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000640

Kines, P. (2002). Construction workers' falls through roofs: Fatal versus serious injuries. *Journal of Safety Research*, *33*(2), 195–208. https://doi.org/10.1016/s0022-4375(02)00019-1

Koc, K., Ekmekcioğlu, Ö., & Gurgun, A. P. (2023). Developing a National Data-Driven Construction Safety Management Framework with Interpretable Fatal Accident Prediction. *Journal of Construction Engineering and Management*, *149*(4), 04023010. https://doi.org/10.1061/JCEMD4.COENG-12848

Koulinas, G. K., Demesouka, O. E., Marhavilas, P. K., Orfanos, N. I., & Koulouriotis, D. E. (2023). Multicriteria Health and Safety Risk Assessments in Highway Construction Projects. *Sustainability*, *15*(12), Article 12. https://doi.org/10.3390/su15129241

Li, J. Q., & Yu, W. (2021). Enhanced Safety Performance Function for Highway Segments in Oklahoma. *Journal of Infrastructure Systems*, *27*(3), 04021018. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000616

Li, Y., & Bai, Y. (2008). Comparison of characteristics between fatal and injury accidents in the highway construction zones. *Safety Science*, *46*(4), 646–660. https://doi.org/10.1016/j.ssci.2007.06.019

Lipscomb, H. J., Glazner, J. E., Bondy, J., Guarini, K., & Lezotte, D. (2006). Injuries from slips and trips in construction. *Applied Ergonomics*, *37*(3), 267–274. https://doi.org/10.1016/j.apergo.2005.07.008

Liu, Y., Ye, G., Xiang, Q., Yang, J., Miang Goh, Y., & Gan, L. (2023). Antecedents of construction workers' safety cognition: A systematic review. *Safety Science*, *157*, 105923. https://doi.org/10.1016/j.ssci.2022.105923

Loayza Chahuayo, P. (2011). *Safety Issues Among Hispanic Construction Workers Along the Wasatch Front in Utah* [Brigham Young University]. https://scholarsarchive.byu.edu/etd/2862

Lokesha, E. B., Aruna, M., Reddy, S. K., & Srinivasa, A. S. (2023). Development of Regression Model and Optimization of Mechanical Properties of Geopolymer Concrete Prepared Using Gold Ore Tailings. *Journal of Hazardous, Toxic, and Radioactive Waste*, *27*(4), 04023033. https://doi.org/10.1061/JHTRBP.HZENG-1259

Loosemore, M., & Malouf, N. (2019). Safety training and positive safety attitude formation in the Australian construction industry. *Safety Science*, *113*, 233–243. https://doi.org/10.1016/j.ssci.2018.11.029

Lyu, S., Hon, C. K. H., Chan, A. P. C., Jiang, X., & Skitmore, M. (2023). Critical Factors Affecting the Safety Communication of Ethnic Minority Construction Workers. *Journal of Construction Engineering and Management*, *149*(2), 04022173. https://doi.org/10.1061/JCEMD4.COENG-12680

Manzoor, B., Othman, I., & Manzoor, M. (2021). Evaluating the critical safety factors causing accidents in high-rise building projects. *Ain Shams Engineering Journal*, *12*(3), 2485–2492. https://doi.org/10.1016/j.asej.2020.11.025

Marji, L. K., Zech, W. C., & Kirby, J. T. (2023). State-of-the-Practice Survey: United States Departments of Transportation Worker Injuries and Safety Program Efforts. *Safety*, *9*(4), Article 4. https://doi.org/10.3390/safety9040067

McKinnon, R. C. (2012). *Safety Management: Near Miss Identification, Recognition, and Investigation*. CRC Press.

Mehmood, I., Li, H., Qarout, Y., Umer, W., Anwer, S., Wu, H., Hussain, M., & Fordjour Antwi-Afari, M. (2023). Deep learning-based construction equipment operators' mental fatigue classification using wearable EEG sensor data. *Advanced Engineering Informatics*, *56*, 101978. https://doi.org/10.1016/j.aei.2023.101978

Metro, K., Harper, C., & Bogus, S. M. (2021). Factors Affecting Workforce Resilience in Public Transportation Agencies. *Journal of Management in Engineering*, *37*(4), 04021018. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000920

Mohandes, S. R., Durdyev, S., Sadeghi, H., Mahdiyar, A., Hosseini, M. R., Banihashemi, S., & Martek, I. (2022). Towards enhancement in reliability and safety of construction projects: Developing a hybrid multi-dimensional fuzzy-based approach. *Engineering, Construction and Architectural Management*, *30*(6), 2255–2279. https://doi.org/10.1108/ECAM-09-2021-0817

Muzafar, M. (2021). Building information modelling to mitigate the health and safety risks associated with the construction industry: A review. *International Journal of Occupational Safety and Ergonomics*, *27*(4), 1087–1095. https://doi.org/10.1080/10803548.2019.1689719

Namian, M., Albert, A., Zuluaga, C. M., & Behm, M. (2016). Role of Safety Training: Impact on Hazard Recognition and Safety Risk Perception. *Journal of Construction Engineering and Management*, *142*(12), 04016073. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001198

Namian, M., Ghorbani, Z., Taherpour, F., Ghiasvand, E., & Karji, A. (2022). Demystifying the Impact of Age on Safety Performance of Construction Workers: Examining the Mediating Roles of Experience and Fatigue. *Practice Periodical on Structural Design* *and Construction*, *27*(4), 04022038. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000718

Nielsen, K., Ng, K., Vignoli, M., Lorente, L., & Peiró, J. M. (2023). A mixed methods study of the training transfer and outcomes of safety training for low-skilled workers in construction. *Work & Stress*, *37*(2), 127–147. https://doi.org/10.1080/02678373.2022.2086646

Nnaji, C., Jafarnejad, A., & Gambatese, J. (2020). Effects of Wearable Light Systems on Safety of Highway Construction Workers. *Practice Periodical on Structural Design and Construction*, *25*(2), 04020003. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000469

Nnaji, C., Lee, H. W., Karakhan, A., & Gambatese, J. (2018). Developing a Decision-Making Framework to Select Safety Technologies for Highway Construction. *Journal of Construction Engineering and Management*, *144*(4), 04018016. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001466

Nnamoko, N., Arshad, F., England, D., Vora, J., & Norman, J. (2014, June 24). *Evaluation of Filter and Wrapper Methods for Feature Selection in Supervised Machine Learning*. The 15th Annual Postgraduate Symposium on the convergence of Telecommunication, Networking and Broadcasting.

Osman, M., Mishra, S., Paleti, R., & Golias, M. (2019). Impacts of Work Zone Component Areas on Driver Injury Severity. *Journal of Transportation Engineering, Part A: Systems*, *145*(8), 04019032. https://doi.org/10.1061/JTEPBS.0000253

Paap, K. (2006). *Working Construction: Why White Working-class Men Put Themselves--and the Labor Movement--in Harm's Way*. Cornell University Press.

Park, I., Kim, J., Han, S., & Hyun, C. (2020). Analysis of Fatal Accidents and Their Causes in the Korean Construction Industry. *Sustainability*, *12*(8), Article 8. https://doi.org/10.3390/su12083120

Rahman, M., & Kang, M.-W. (2021). Analysis of Intersection Site-Specific Characteristics for Type II Dilemma Zone Determination. *Journal of Transportation Engineering, Part A: Systems*, *147*(10), 04021072. https://doi.org/10.1061/JTEPBS.0000578

Rokooei, S., Shojaei, A., Alvanchi, A., Azad, R., & Didehvar, N. (2023). Virtual reality application for construction safety training. *Safety Science*, *157*, 105925. https://doi.org/10.1016/j.ssci.2022.105925

Sabeti, S., Shoghli, O., Baharani, M., & Tabkhi, H. (2021). Toward AI-enabled augmented reality to enhance the safety of highway work zones: Feasibility, requirements, and challenges. *Advanced Engineering Informatics*, *50*, 101429. https://doi.org/10.1016/j.aei.2021.101429

Saffer, H., Dave, D., Grossman, M., & Leung, L. A. (2013). Racial, Ethnic, and Gender Differences in Physical Activity. *Journal of Human Capital*, *7*(4), 378–410. https://doi.org/10.1086/671200

Salama, D. M., & El-Gohary, N. M. (2016). Semantic Text Classification for Supporting Automated Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, *30*(1), 04014106. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000301

Sanni, O., Ikoba, N., & Adegboye, O. (2021). *Modified Methods of Computing Some Descriptive Statistics for Grouped Data*. *4*, 114–134.

Seo, W., Kim, B., Bang, S., & Kang, Y. (2024). Identifying Key Financial Variables Predicting the Financial Performance of Construction Companies. *Journal of Construction Engineering and Management*, *150*(3), 04024007. https://doi.org/10.1061/JCEMD4.COENG-13959

Smith, C. K., Wuellner, S., & Marcum, J. (2023). Racial and ethnic disparities in workers' compensation claims rates. *PLOS ONE*, *18*(1), e0280307. https://doi.org/10.1371/journal.pone.0280307

Song, P., & Kroll, C. (2012). *The Impact of Multicollinearity on Small Sample Hydrologic Regional Regression*. 3713–3722. https://doi.org/10.1061/41173(414)389

Spangenberg, S., Hannerz, H., & Tüchsen, F. (2005). Hospitalized injuries among bridge and tunnel construction workers. *Construction Management and Economics*, *23*(3), 237–240. https://doi.org/10.1080/01446190500041503

Suraji, A., Duff, A. R., & Peckitt, S. J. (2001). Development of Causal Model of Construction Accident Causation. *Journal of Construction Engineering and Management*, *127*(4), 337–344. https://doi.org/10.1061/(ASCE)0733-9364(2001)127:4(337)

Thakkar, A., & Lohiya, R. (2021). Attack classification using feature selection techniques: A comparative study. *Journal of Ambient Intelligence and Humanized Computing*, *12*(1), 1249–1266. https://doi.org/10.1007/s12652-020-02167-9

Tong, B., Guo, J., & Fang, S. (2021). Predicting Budgetary Estimate of Highway Construction Projects in China Based on GRA-LASSO. *Journal of Management in Engineering*, *37*(3), 04021012. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000890

Tong, R., Zhao, H., Zhang, N., Li, H., Wang, X., & Yang, H. (2020). Modified accident causation model for highway construction accidents (ACM-HC). *Engineering, Construction and Architectural Management*, *28*(9), 2592–2609. https://doi.org/10.1108/ECAM-07-2020-0530

Tork, N. (2008). *The Effects of Gendered Stressors on Female Construction Workers' Occupational Health and Safety: A Re-Analysis of Qualitative Data* [University of Cincinnati].

https://etd.ohiolink.edu/acprod/odb_etd/etd/r/1501/10?clear=10&p10_accession_num=ucin1212071613

Tözer, K. D., Gürcanli, G. E., Çelik, T., & Akboğa Kale, Ö. (2024). Safer designs with BIM based fall hazards identification and accident prevention. *Safety Science*, *169*, 106346. https://doi.org/10.1016/j.ssci.2023.106346

U.S. Department of Transportation. (2024). *Budget Highlights 2024*. https://www.transportation.gov/sites/dot.gov/files/2023-03/BudgetHL2024_Mar09_3pm_508.pdf

U.S. DOT. (2023a). *Bipartisan Infrastructure Law (BIL) / Infrastructure Investment and Jobs Act (IIJA) | PHMSA*. https://www.phmsa.dot.gov/legislative-mandates/bipartisan-infrastructure-law-bil-infrastructure-investment-and-jobs-act-iija

U.S. DOT. (2023b). *Performance Plan and report*. https://www.transportation.gov/sites/dot.gov/files/2023-08/FY_2024_APP-FY_2022_APR_508_Compliant_5.10.pdf

Valcamonico, D., Baraldi, P., Amigoni, F., & Zio, E. (2022). A framework based on Natural Language Processing and Machine Learning for the classification of the severity of road accidents from reports. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 1748006X221140196. https://doi.org/10.1177/1748006X221140196

Wang, D., Dai, F., & Ning, X. (2015). Risk Assessment of Work-Related Musculoskeletal Disorders in Construction: State-of-the-Art Review. *Journal of Construction Engineering and Management*, *141*(6), 04015008. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000979

Wang, X., Dong, X. S., Choi, S. D., & Dement, J. (2017). Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014. *Occupational and Environmental Medicine*, *74*(5), 374–380. https://doi.org/10.1136/oemed-2016-103943

Weng, J., Zhu, J.-Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis & Prevention*, *92*, 43–52. https://doi.org/10.1016/j.aap.2016.03.017

Williams, G. (2011). Descriptive and Predictive Analytics. In G. Williams (Ed.), *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery* (pp. 171–177). Springer. https://doi.org/10.1007/978-1-4419-9890-3_8

Wong, L. S., Marani, A., & Nehdi, M. L. (2021). Gradient Boosting Coupled with Oversampling Model for Prediction of Concrete Pipe-Joint Infiltration Using Designwise Data Set. *Journal of Pipeline Systems Engineering and Practice*, *12*(3), 04021015. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000557

Ye, Z., Ye, Y., Zhang, C., Zhang, Z., Li, W., Wang, X., Wang, L., & Wang, L. (2023). A digital twin approach for tunnel construction safety early warning and management. *Computers in Industry*, *144*, 103783. https://doi.org/10.1016/j.compind.2022.103783

Zermane, A., Mohd Tohir, M. Z., Zermane, H., Baharudin, M. R., & Mohamed Yusoff, H. (2023). Predicting fatal fall from heights accidents using random forest classification machine learning model. *Safety Science*, *159*, 106023. https://doi.org/10.1016/j.ssci.2022.106023

Zhang, K., & Hassan, M. (2019). Identifying the Factors Contributing to Injury Severity in Work Zone Rear-End Crashes. *Journal of Advanced Transportation*, *2019*, e4126102. https://doi.org/10.1155/2019/4126102

Zhang, W., & Zhang, S. (2020). Chicago Traffic Collision Data Analysis Based on Multi-Component Analysis and Exploratory Data Analysis. 4684–4696. https://doi.org/10.1061/9780784483053.389

Zhang, Y., Minchin, R. E., & Agdas, D. (2017). Forecasting Completed Cost of Highway Construction Projects Using LASSO Regularized Regression. Journal of Construction Engineering and Management, 143(10), 04017071. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001378

Zhou, X., & Liao, P.-C. (2023). Weighing Votes in Human–Machine Collaboration for Hazard Recognition: Inferring a Hazard-Based Perceptual Threshold and Decision Confidence from Electroencephalogram Wavelets. Journal of Construction Engineering and Management, 149(9), 04023084. https://doi.org/10.1061/JCEMD4.COENG-13351