Teknik Dergi, 2020 9919-9940, Paper 573

Identifying Factors that Contribute to Severity of Construction Injuries using Logistic Regression Model

Özge AKBOĞA KALE[1]
Selim BARADAN[2]

ABSTRACT

Majority of research in occupational safety and health area lean towards describing accidents with the aid of surveys and descriptive statistics, instead of using inferential statistical techniques. Therefore, an extensive archival study was performed in cooperation with Social Security Institute of Turkey, which included examination and reorganization of more than 2000 accident report forms to create a categorically identified data set, incorporating "Injury Severity Score" concept, followed by various statistical analysis techniques (univariate frequency, cross tabulation and binary logistic regression). As a result, a model was developed to identify the factors that contribute to severity. The findings of the analyses showed that four of the independent variables (work experience, accident type, unsafe condition and unsafe act) have statistically significant influence on workplace injury severity.

Keywords: Occupational safety and health, logistic regression analysis, injury severity score, construction accidents, data mining.

1. INTRODUCTION

Research efforts in occupational safety and health field have mainly concentrated on determining the causes of occupational injuries and illnesses and thus help discover new strategies to reduce or eliminate them. Majority of research studies in the past used surveys to collect information from employers and workers [1-3], while several studies preferred to use descriptive statistics to summarize and interpret occupational injury and illness data [4-12]. Even though conducting surveys and using descriptive statistics may be helpful, such as describing how the accident occurred and who were involved; using them alone is not sufficient in determining the factors that contribute most to the unfortunate incident. Therefore, using more advanced statistics techniques, such as logistic regression, to support

---

Note:

- This paper has been received on October 15, 2018 and accepted for publication by the Editorial Board on January 28, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.470633

1 İzmir Demokrasi University, Department of Civil Engineering, Izmir, Turkey - ozge.akbogakale@idu.edu.tr - https://orcid.org/0000-0002-3848-0578

2 Ege University, Department of Civil Engineering, Izmir, Turkey- selim.baradan@ege.edu.tr - https://orcid.org/0000-0002-9172-8552

findings coming from surveys and descriptive statistics could be invaluable in accident causation studies.

Logistic regression, which is a commonly used inferential statistics technique in medical and social science to develop statistical models and make predictions [13-17], is rarely utilized in occupational safety and health research [18-20], most likely, this study aims to fill this gap in occupational safety and health field by integrating inferential statistics, namely logistic regression technique, to develop a statistical model in identifying the factors that contribute most to the severity of occupational injuries. The study also benefited from traditional descriptive statistics techniques to analyze frequency distribution, which are discussed in the following sections.

Severity is particularly an important topic to focus on. Risk is typically defined as the combination of two components: the probability (frequency) of a defined hazard and the consequences of its occurrence (severity). In occupational safety and health field, severity is more difficult to estimate when compared to frequency. Frequency can be merely calculated by finding the ratio between number of incidents and total number of workers or loss of time due to injury and total work hours. On the other hand, severity is the potential loss when an event occurs and the loss may be expressed in human terms, such as loss of life, serious injury, serious illness, number of losses and so forth [21]. Therefore, efforts were made in this study to find methods to properly quantify and analyze severity component of occupational safety and health risk.

The reason for the drought of advanced statistical studies in occupational safety and health research is most likely due to available public data being ambiguous and insufficient. Similarly, a detailed database was paramount for the intended study to proceed. Thus, cooperation was established with Social Security Institution (SSI), which collects injury statistics in Turkey, to access their database for archival study. Turkish construction industry was used as the sample data set, since construction is the leading industry in fatal injuries (35.3%) in Turkey, and still remains a major problem with ever increasing construction projects and catastrophic incidents along with them [22-24], Consequently, this study also aims to depict an accurate safety portrait of Turkish construction industry, taking into account that there are not many statistical studies due to improper recordkeeping system and lack of immediate data in construction.

The objectives of the research study addressed in this paper can be summarized as: (a) identifying the factors affecting the injury severity score of construction injuries; (b) investigating the frequency distributions of these factors; (c) examining the relationships between the factors; and (d) developing a predictive statistical model for construction injuries in Turkey.

2. MATERIALS AND METHODS

The research study summarized in this paper consisted of two parts: The first part was mainly the creation of database to be used in statistical analyses. Physical reports of more than two thousand construction accidents that resulted in fatal or nonfatal injury were collected from the offices of Social Security Institution in three different cities (İstanbul, Ankara, İzmir), since there was not an available online database. Then the raw data in the reports were organized to generate a new database to be used for statistical analyses. The database included 16 research variables, where each variable had several categories. Injury severity score was selected as the dependent variable, while the others were assumed as independent variables.

The second part of the study was applying a methodology that included performing three different statistical analysis techniques on the database. First, descriptive univariate frequency analysis was performed to establish data demographics and observe the distribution of selected variables with using frequency tables. Then in the next step, cross-tabulation, a bivariate analysis method, was used to examine the relationships between each independent variable and the dependent variable. The final statistical analysis in the study was applying binary logistic regression on the identified independent variables to come up with a model that predicts the injury severity. The flow chart of the methodology is displayed in Figure 1.

The most challenging parts of the methodology were to quantify injury severity for logistic regression, where many factors affect the outcome, and how to integrate fatal and nonfatal injuries. Therefore, a literature review was conducted to find a universally accepted method to quantify injury severity. Even though the review failed to find such a method used in construction safety and health area, a rating system named "Injury Severity Score (ISS)", which is predominantly seen in medicine and traffic safety research studies, was adopted for this study. Using ISS to quantify severity of construction injuries enabled ranking and making comparisons between each injury case. Furthermore, ISS was a great fit for logistic regression analysis part of the study due to ISS having a nonlinear structure.

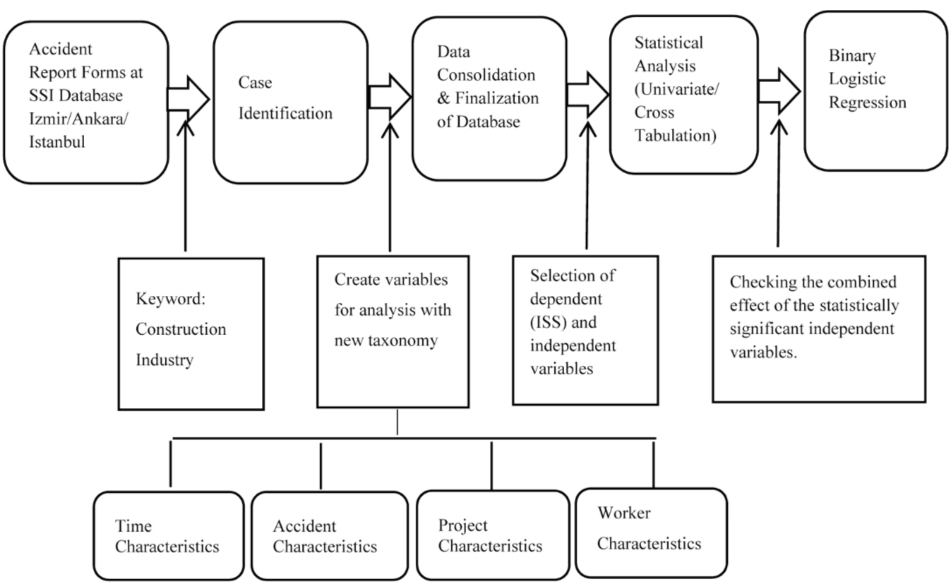

Figure 1 - Flow chart of the methodology

3. DATA ACQUISITION AND MINING

Finding occupational injury data is a significant challenge in Turkey, even though employers are responsible for reporting worksite injuries and illnesses to Social Security Institution (SSI) as mandated by law (Article 14) [25]. SSI shares annual summaries of these reports on their website [22]. However, it isn't possible to extract detailed information from the available online data. Therefore, a special work permit was obtained to be able to study injury reports stored in SSI's physical archives in three biggest cities of Turkey (Ankara, Istanbul, and Izmir).

Following this arrangement, a systematic archival study was carried out. Since, work-related injury reports lack an industry or injury based categorization system, report forms associated with construction injuries had to be handpicked among 50,000 work-related accident cases. This data collection effort covered only the cases reported between 2010 and 2012. During data collection process there have been no important changes in the way that accident data were collected or in compensation practices. The classification system of accidents and occupations remained unchanged as well.

The selection process included identifying fatal and nonfatal injuries that only occurred in the construction sites and production areas, such as ready-mix concrete, prefabricated structural concrete and steel production. Accident cases in other industries were not taken into consideration. Overall 2249 construction related cases were found in the database (1062 cases from Izmir, 732 cases from Istanbul, and 655 cases from Ankara) of those, 425 records were disregarded due to inadequate or missing information. Finally, a total of 2024 cases were selected for analyses.

The SSI workplace accident report forms already contain some of the information about the variables within the defined characteristics, such as injured worker characteristics, nature of injury, injured body part and equipment involved. However, existing information wasn't sufficient to perform the intended analysis in this study. Particularly, there were too many inconsistent answers due to the nature of open ended questions contained in the form. Therefore, a new taxonomy was created to be used for descriptive statistical analysis. During this process, a thorough screening of collected data was performed to identify missing data, consolidate some of the categories and include additional variables.

The significance of the pattern of missing data in a data set is indicated by Tabachnik and Fidell (2013) [26]. Deleting the variables or dropping the cases with missing data from the data set are two ways to deal with missing data. Dismissing or dropping the case from the data set is the first alternative if a case is missing too many data. However, just deleting the variable is suggested if only certain variable information is missing for too many cases. The data set of this study had only one problematic variable in terms of missing data; worker's education background. Apart from this, there were also a few unknown variables in several cases. However, all the cases were kept to acknowledge their presence in the database. Only the missing data from certain variables were neglected and weren't included in the analysis.

Additional categorical variables were also required to be added to the new database by using the information extracted from open-ended questions. Kass (1980) also suggests merging some levels in order to reach a meaningful conclusion in studies that involve data mining [27]. Therefore, collapsing levels technique was applied to variables with high number of categories but low number of observation counts.

Finally, taxonomy study resulted in 15 independent research variables. They were grouped under four different categories (time, project, accident, and worker characteristics) according to their relevance with their characteristics, as seen in Tables 1-4. Also, Injury Severity Score was selected as the dependent variable required for bivariate analysis section of the study, which will be addressed in the following section. Independent variables selected for this research were, in most part, consistent with those used in past research [28-33]. More information about the categories and variables listed under them are given below.

Time characteristics describe when the accident happened and include three variables: month of the year, day of the week and hour of the day.

Table 1 - Distribution of Time Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Percentage (%)
Month of accident	January	205	10.1	**Hour of Accident**	07.00<t≤08.00	19	1.1
	February	207	10.2		08.00<t≤09.00	152	8.5
	March	208	10.3		09.00<t≤10.00	216	12.0
	Aprıl	181	8.9		10.00<t≤11.00	225	12.5
	May	171	8.4		11.00<t≤12.00	158	8.8
	June	153	7.6		12.00<t≤13.00	72	4.0
	July	166	8.2		13.00<t≤14.00	174	9.7
	August	117	5.8		14.00<t≤15.00	189	10.5
	September	132	6.5		15.00<t≤16.00	196	10.9
	October	182	9.0		16.00<t≤17.00	157	8.8
	November	112	5.5		17.00<t≤18.00	81	4.5
	December	190	9.4		18.00<t≤19.00	40	2.2
Day of Accident	Monday	349	17.2		19.00<t≤20.00	23	1.3
	Tuesday	296	14.6		20.00<t≤07.00	92	5.1
	Wednesday	320	15.8				
	Thursday	313	15.5				
	Friday	305	15.1				
	Saturday	253	12.5				
	Sunday	188	9.3				

Project characteristics provide information about the construction project that the worker was involved in (type and end use) at the time of incident. Construction sites, different in shape and size, are unique dynamic environments. Therefore, project characteristics help to understand and classify the construction environment where accidents mostly occur. Project type variable comprises of 3 selections: new project or addition, manufacturing of construction materials, and repair/maintenance/renovation of existing structures. Project end use variable addresses the usage purpose of the finalized project (residential, institutional and commercial, industrial, infrastructure and construction materials).

Table 2 - Distribution of Project Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Valid Percentage (%)
Project Type	New project or new addition	1710	84.5	Project End Use	Residential	1334	66.3
	Manufacturing of construction materials	255	12.6		Construction Materials	255	12.7
	Repair/Maintenance/ Renovation	59	2.9		Institutional and Commercial	243	12.1
					Infrastructure / Heavy Construction	167	8.3
					Industrial	13	0.6

Accident characteristics variables (type of injury, nature of injury, injured body part, unsafe act and condition) reveal plenty of information regarding the incident and injury; in other words, they describe the accident. Categorical data under “type of injury” variable were created based on the International Classification of Diseases (ICD) codes [34], while nature of injury and injured body part information were directly taken from workplace accident reports. Unsafe acts and conditions variables were deducted from the open ended questions in the workplace accident report forms based on the ILO accident cause theory and Hill’s definitions [35-36]. Heinrich's Domino Theory states that accidents result from a chain of sequential events, metaphorically like a line of dominoes falling over. According to Herbert W. Heinrich; all incidents directly relate to unsafe conditions and acts and removing a key factor (an unsafe condition or an unsafe act) prevents the start of the chain reaction [37].

Worker characteristics variables aim to investigate the personal (Age, education) and professional (work experience, assigned task, construction trade) background of injured worker. Data categories were mostly determined in compliance with the classification used in SSI annuals.

Table 3 - Distribution of Accident Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Percentage (%)
Injured Body part	Upper extremity	623	32.7	Nature of Injury	Superficial wound	496	26.2
	Lower extremity	444	23.3		Bruise	379	20.0
	head	359	18.9		Fracture/Crack	262	13.8
	Multi-injury	218	11.5		Cut	235	12.4
	Whole body	118	6.2		Skeletal and muscular system disorders	132	7.0
	Internal organs	81	4.3		Death	99	5.2
	Back	50	2.6		Foreign object in the eye	91	4.8
	Neck	10	0.5		Strain	75	4.0
Unsafe condition	Insufficient/Lack of Written Work Procedures	640	32.1		Trauma and Internal injury	67	3.5
	Poor Housekeeping	484	24.3		Electrocution	26	1.4
	No collective protection systems	387	19.4		Burns (heat or chemical)	24	1.2
	No PPE provided	293	14.7		Others	10	0.5
	Faulty tool/equipment/machinery	191	9.5	Type of injury	Falls	734	36.4
Unsafe Act	No PPE Usage	678	34.0		Struck by falling object	604	30.0
	Unsafe work practices	604	30.3		Bite/Sting/Scratch	283	14.1
	Position Inappropriate for Task	312	15.6		Caught In or Between	260	12.9
	Unsafe act by a third party	222	11.1		Others	52	2.6
	Defective/Inappropriate Equipment In Use	179	9.0		Electric shock	29	1.4
					Vehicle accident	24	1.2
					Fire / Explosion	18	0.9
					Heat exposure	10	0.5

Table 4 - Distribution of Worker Characteristics

Variable	Description	Frequency Count	Frequency Percentage (%)	Variable	Description	Frequency Count	Frequency Percentage (%)
Worker's age	14-15	1	0.1	Assigned task	Formwork	494	25.3
	16-18	36	2.0		Material handling	342	17.5
	19-24	373	20.3		Plaster/Paint	174	8.9
	25-29	362	19.7		Assembly/disassembly	168	8.6
	30-34	328	17.9		Ironwork	155	7.9
	35-39	245	13.4		Concrete handling	118	6.0
	40-44	211	11.5		Maintenance / Repair	98	5.0
	45-49	176	9.6		Other services	91	4.7
	50-54	65	3.5		Commuting	58	3.0
	55-59	32	1.7		Field Inspection	54	2.8
	60-64	3	0.2		Woodwork	44	2.3
	65+	1	0.1		Welding	42	2.1
Work experience in project	Beginning at same day	167	8.6		Work break	35	1.8
	1. Week	170	8.7		Installation	23	1.2
	2-4 Weeks	455	23.4		Excavation work	20	1.0
	1-3 Months	551	28.3		Sheet metal work	18	0.9
	3-6 Months	289	14.8		Marble and tile setting	16	0.8
	6-12 Months	166	8.5		Asphalt paving	5	0.2
	12-24 Months	40	2.1	Construction trade	Unskilled worker	563	28.1
	>24 Months	110	5.6		Form worker	512	25.5
Educational Background	Literate	93	8.8		Ironworker	181	9.0
	Illiterate	21	2.0		Assembly/Installation/Maintenance	162	8.1
	Elementary school	567	53.7		Plasterer/Painter	159	7.9
	Middle school	317	30.0		Heavy equipment operator	112	5.6
	High school	29	2.7		Foreman	97	4.8
	Vocational high school	11	1.0		Welder	64	3.2
	Undergraduate	16	1.5		Plumber	40	2.0
	Graduate	1	0.1		Technician	36	1.8
					Installation	26	1.3
					Tile setter	17	0.8
					Security guard	14	0.7
					Electrician	13	0.6
					Carpenter	10	0.4

4. DATA ANALYSIS

4.1. Injury Severity Score (ISS)

Appropriate classification of injuries by type and severity is fundamental to the study of injury [38]. Even within a single community, groups of injured persons differ as to the nature and severity of their injuries. The difficulty of adjusting for such variation has hampered scientific study of injured persons. Nevertheless, it is essential to take differences in severity of injury into account when comparing the morbidity (the state of being diseased or unhealthy) and mortality (the number of people who died) of various groups for purposes of evaluating their emergency and subsequent care [39].

Quantifying injury severity has been a challenge in this study as well. A literature review was conducted to find a universally accepted method to quantify injury severity. The Injury Severity Score (ISS), a method predominantly seen in medicine and traffic safety research studies, was found. ISS is virtually the only anatomical scoring system in use, and correlates linearly with mortality, morbidity, hospital stay and other measures of severity [40]. Therefore, this study adapted ISS rating system to quantify severity of construction injuries, and to rank and make comparisons between each injury case. It was also deemed a great fit for logistic regression analysis part of the study due to its nonlinear structure.

Table 5 - Classification of ISS

ISS	Consequence	Symptom	Count	Percentage (%)
≤ 3	First aid was not needed. No time away from work.	Strain, skeletal and muscular system disorders	441	23.3
4 ≤ ISS ≤ 8	First aid was given but no medical intervention was needed. No long time away from work.	Superficial wound, bruise	743	39.2
9≤ ISS ≤ 24	Medical intervention is needed. After long time there is a return to work.	Cut, fracture, trauma and internal injury	539	28.4
≥ 25	Mostly workforce loss or life loss.	Electrocution, multi-fracture, death	173	9.1

The ISS can be applied to persons who have sustained injury to more than one area of the body as well as to those with isolated injuries. Each injury is assigned an Abbreviated Injury Score (AIS) (may vary from 1-minimum severity to 6-maximum severity, almost always fatal) and is allocated to one of six body regions (head, face, abdomen, extremities, external). Only the highest AIS in each body region is used. An individual's ISS is determined by rating each injury with the AIS, then adding together the squares of the highest AIS rating for each of the three most severely injured body areas [40]. The ISS takes values from 0 to 75; it can generate only 44 values [42-44]. The ISS gives equal importance to injuries with the same AIS severity occurring in different body regions [45]. Researchers used ISS for different areas such as medicine, traffic safety [46-51]. Different categorizations were used for different cases and application areas in these studies. It is noteworthy that categorization was done in a systematic way but there was no standardization. With this in mind, the ISS calculated for the cases in this study was classified within a logical framework. ISS for each victim was estimated by using the type of injury and injured body part information found in the SSI workplace accident report forms. ISS for each casualty, ranging from 0 to 75, was calculated and then classified into 4 groups. Classification was done based on type of injury and time away from work. The framework used for ISS classifications along with the counts and percentages of analyzed cases are displayed in Table 5.

ISS, was used as a dependent quantitative variable in addition to the 15 variables grouped under four categories that are presented in Tables 1-4. ISS variable was converted into binary category at the modelling stage in the logistic regression analysis. Converting process is explained in detail in the Logistic Regression Analysis section of this paper.

4.2. Univariate Analysis

In vast majority of the construction safety literature that utilizes descriptive statistics analysis, research findings are based on univariate analysis [52-55]. Similarly, univariate analysis was adopted in this research, for two purposes: data screening and classification, and to understand what we have and choose the right variables for bivariate data analysis, which is the next step in the methodology.

As discussed in the previous section of this paper, the database created for this study consists of 15 research variables. The frequency distributions of each variable are presented in Tables 1-4. Each table and corresponding variables were analyzed and interpreted as part of univariate analysis. Results of this analysis and key findings are presented in Results and Discussion section.

4.3. Cross Tabulation Analysis

After applying univariate analysis to investigate whether there is a meaningful relationship between variable pairs, cross tabulation analysis is carried out. Cross tabulation is widely utilized to study the relationship between categorical variables [55]. Researchers used cross tabulation analysis for different purposes, such as for determination of the relationship between service quality and customer satisfaction, comparing three widely used methods for assessing alcohol consumption, tracing of the pattern of general and specific aspects of marital satisfaction over the family life cycle and determining attitudes of students towards using credit cards [56-60].

Cross tabulation analysis produces a contingency table displaying relationship, in the form of joint frequencies of two or more variables. The rows indicate one variable while the columns indicate the other [61]. The Pearson Chi-square test is one of the tests that can be used to interpret this relationship successfully [62]. Pearson chi-square compares the observed counts with those that would be expected if there were no association between two variables [63]. After calculating the Pearson Chi-squared value, the p-value based on that value (which expresses the importance of the Chi-square value) must be calculated separately. The p-value is the probability value that is used for hypothesis testing by the Pearson chi-square test. After finding the p-value, one can decide whether the result is statistically significant or not. Most common practice for significance level is 0.05, in other words the confidence interval is 95%. Therefore, a p-value less than 0.05 is accepted as significant and allows researcher to reject the null hypothesis of no association, and conclude that there is an association between variables [62].

After the null hypothesis is rejected to determine the strength of this relationship, researchers have to calculate Phi or Cramer's V values. Phi values from 0 to 0.1 show a weak relationship; 0.1 to 0.3 indicate a moderate relationship; and values between 0.3 and 1.0 suggest a strong relationship [64].

Following the conclusion of univariate frequency analysis in this study, Injury severity score (ISS) was selected as the dependent variable, while the others were assumed as independent variables. Cross tabulation analysis was performed between dependent variable ISS and each independent variable. Pearson chi-square, "p" and Phi or Cramer's V values were taken into account to statistically test the relationships of variables.

Even though, cross tabulation analysis investigates whether a significant relationship exists between the dependent and independent variables and its influence on the statistically dependent variable, the findings of the cross tabulation do not reveal information about the interactions between the independent variables. Therefore, binary logistic regression analysis was performed to check the combined effect of the statistically significant independent variables and to determine their relationships with the dependent variable.

4.4. Logistic Regression Analysis

Logistic regression (LR) is a statistical analysis tool for modeling the relationship between a response variable and a set of explanatory variables when the response variable is categorical [26]. The main goal of the LR analysis is to find the best fitting and most parsimonious, yet reasonable, model to describe the relationship between a dependent (response) and a set of independent (predictor or explanatory) variables [65]. Logistic regression does not have the requirements for the independent variables to be normally distributed, linearly related or of equal variance within each group [67]. In logistic regression, instead of predicting the value of Y (dependent variable) from predictor variable X1…n , the probability of Y occurring is predicted with given known values of $X_{1\ldots n}$ [61].

The significance of LR lies in the logistic transformation. In order to predict the dependent variable probability and, perform logistic transformation, the probability function can be written as Equation 1, where p is the probability of being in one group (occurrence of an event) and 1-p is the probability of being in the other group (non-occurrence of an event) [67].

$$\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \qquad (1)$$

The problem with Equation 1 is that the right side of the equation can get any value between - ∞ to + ∞ while, the left side of the equation cannot be negative. The logit transformation equation must be used to overcome this problem, and it is formulated as the following equation where the natural log of the probability of being in one group divided by the probability of being in the other group.

$$logit\ (p) = \ln\left(\frac{p}{1-p}\right) \tag{2}$$

After logit transformation is applied the equation becomes;

$$\ln\left(\frac{p}{1-p}\right) = \ln(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon) \tag{3}$$

In solving the Equation 3, the LR equation from which the probability of Y is predicted becomes as the following equation where, P(Y) = probability of Y occurring; e is the base of natural logarithm and β_0 represents exposure variable or constant, $\beta_{1...n}$ are the coefficients, and $X_{1...n}$ are the independent (predictor) variables.

$$P(Y) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n+\varepsilon)}} \tag{4}$$

In LR, to distinguish the difference between the occurrence and non-occurrence of an event, the dependent variable is coded in a certain way. Assigning a value of 1 (Y=1) to event occurrence and 0 (Y=0) to no occurrence is the simplest way to code the dependent variable [26]. In this research, injury severity score (ISS) was selected as a binary dependent variable. A value of one (1) was assigned to the variable, when ISS is ≥ 9; otherwise, it was assigned zero (0) indicating an ISS < 9. Independent variables identified as statistically significant in cross tabulation analysis were treated as multilevel independent variables.

In order to measure the performance of model's validation of the LR models is necessary.

Poorly fitting results that inaccurately predict the future outcomes may be caused doesn't apply validation to the model [68]. In this study, the model subsets were selected based on a 70/30 ratio. Bernoulli distribution was used to facilitate a random selection on injury cases. Hence Bernoulli distribution takes values of 0 and 1, 1 value was assigned randomly to 70% of the cases, which were used to develop the model and the remaining 30% was used to validate the data [69].

Initially, prediction power of the naïve model, which has only one constant (β_0) without any predictor variables, was estimated. Then, independent variables were added to the naïve model to improve its predictive power, where the improved model is called the development model in this study. A stepwise backward enter method was used during the development of the model. After incorporating the independent variables into the model, the variables that did not have significant positive effect on the predictive power of the development model were left out (assigned task, construction trade), while the significant ones remained in the model (work experience in project, type of injury, unsafe condition, and unsafe act). It should be noted that validation of 30% data set was performed for both naïve and development models.

The model creation is a challenge, to choose the best predictive model various numbers of tests should be applied. The first thing one has to do is to make sure that it meets the guidelines for "goodness-of-fit". This goodness-of fit is done by a parameter that checks the fit of the model. The log-likelihood needs to be calculated In order to do so. The log-likelihood is based on summing the probabilities associated with the predicted and actual outcomes [26]. Wald's test, Hosmer and Lemeshow's R_L and Exp (β) can be listed as other tests that need to be conducted. Wald test is used to determine whether an independent variable is a significant predictor of the outcome or not. Hosmer and Lemeshow's R_L is a test which represents the measure of how much the goodness of fit improves as a result of the inclusion of predictor variables in each step. Exp(β) is the exponential value of the β coefficients and its value represents the odds ratio. Therefore, Exp(β) represents the odds ratio of that predictor variable and how it affects the outcome [61]. Similar to the past research discussed above, this study calculated log-likelihood values, ran Wald and Hoshmer Lemeshow chi-squared tests to determine the predictive power of the development model. Interpretation of the results can be found in Results and Discussion sections.

5. RESULTS

5.1. Univariate Analysis Results

The results of univariate analysis are presented in Table 1-4, where the "count" values of every variable are displayed. This value indicates how many cases are recorded in each category, or in other words their frequency. "Percentage" value in the table represents the percentage distribution of the data categorized under the same variable type. The following inferences were drawn by analyzing the frequency distribution of each variable:

- Contrary to expectations, the number of accidents in winter turned out to be higher than in summer. This finding is quite surprising because the production output in construction industry, which is negatively affected by weather conditions in winter, usually accelerates during summer months. It was expected to see more injuries in summer due to increase in the number of workers on site. This could indicate that there are fewer risks in construction sites, while working under fair weather conditions.
- It was seen that injuries occurred on Monday the most (17.2%). Many previous research studies have proved that this finding is common, which is also named as "Monday Effect" [67-70].
- It was observed that the injuries tend to occur more between the hours of 10:00-12.00 (24.5%) and 14.00-16:00 (21.4%). It is assumed that this could happen due to acceleration of work during those time periods.
- When "type of injury" variable was examined, "falls" and "struck by falling object" were the two most noticeable data with 36.4% and 30.0% frequencies respectively. These findings show similarities with the past studies. Many researches pointed out that fall accidents are not only the highest frequent in the construction but also their results are more severe [71-76]. Similarly, Arndt et al., found that struck by falling objects is one of the common causes of fatal injuries in the construction industry [77].
- There wasn't a clear cut choice among the results of "nature of injury" variable category. Superficial wound, bruise, fracture/crack and cuts were the most encountered injuries. However, these results did not lead to a specific finding.
- Findings indicate that workers who are assigned in formwork and material handling (such as lifting and carrying materials and products) had the most number of injuries with 25.3%

and 17.5% respectively. Thus, these two construction tasks pose more risks to workers safety according to this study.

- It is observed that the worker's age group is predominantly younger. The average age of victims in the database is 33 and the workers between 19-24 years old (20.3%) had more injuries than the other age groups. It was found that the frequency of accidents decreased in proportion to the increasing age.
- Results of the study showed that work experience reduces injury severity. This somewhat anticipated finding solidifies that the increasing work experience has a positive effect on safety awareness. As the worker gains experience, they become experts in their jobs and become more aware of their surroundings with safety in mind. Safety and health training, if provided, also is a tool for reducing injury severity. Past research on similar topic also could support this finding [11].
- Educational background variable revealed an interesting finding that in more than half of the cases analyzed, workers had no further education beyond elementary school. This finding is directly associated with the fact that majority of construction force in Turkey has elementary school diploma or less.
- Similar to educational background finding, most of the cases (28.1%) involved unskilled worker because most of the construction workforce consists of unskilled laborers. However, a noteworthy amount (25.5%) of the cases involved form workers. This finding naturally matches with the findings of "assigned task" category, which is somewhat close to the description of "construction trade" variable.
- When unsafe act and condition variables were analyzed, three data categories stood out: no personal protective equipment (PPE) usage, insufficient/lack of written work procedures and unsafe work practices. Other most encountered unsafe acts and conditions were poor housekeeping, no collective protection systems, no PPE provided, and position inappropriate for the task.

5.2. Cross Tabulation Results

To investigate the relationship between ISS and other independent variables cross tabulation analysis was performed as the second part of the descriptive analysis. Only statistically significant results of cross tabulation analysis were presented in Table 6. According to the analysis, 6 of the 14 independent variables (work experience in project, assigned task, construction trade, type of injury, unsafe act, unsafe condition) were found statistically significant with $p<0.05$, as shown in Table 6.

Table 6 also displays Cramer's V values, which are used for determining the strength of relationship between variables. Findings show that work experience in project (Crv: 0.142), assigned task (Crv: 0.153), construction trade, (Crv: 0.139) and unsafe act (Crv: 0.245) have moderate relationship strength, while type of injury (Crv: 0.347) and unsafe condition (Crv: 0.330) has a strong relationship with the dependent variable (Table 6).

Table 6 - Contingency Table - ISS vs. Nominal Variables

Independent Variables	Pearson's X^2(df), p		Phi Cramer's V	
Work experience in project	$X^2(6)=27,676$	p=0.000	crv(6)=0.1420	p=0.000
Assigned task	$X^2(12)=31,112$	p=0.001	crv(12)=0.153	p=0.001
Construction trade	$X^2(7)=26,705$	p=0.000	crv(7)=0.139	p=0.000
Type of injury	$X^2(8)=166,001$	p=0.000	crv(8)=0.347	p=0.000
Unsafe Condition	$X^2(4)=149,508$	p=0.000	crv(4)=0.330	p=0.000
Unsafe Act	$X^2(4)=82,666$	p=0.000	crv(4)=0.245	p=0.000

5.3. Logistic Regression Analysis Results

Naïve and development models were generated, validated and tested for prediction power in the logistic regression analysis part of the study. Injury Severity Score (ISS) was used as the dependent variable for both models. While naïve model used six variables that were identified in cross tabulation analysis part, development model left out two variables (assigned task, construction trade) that did not have significant positive effect on the predictive power of the model. The following results were obtained from the analysis:

The naïve model shows the general percentages of the classified data without any predictor variables and uses the 70% data set, which is assumed to have ISS lower than 9. The prediction power of the naïve model was found to be 61.5%. Whereas the prediction power of the 30% data set, which is used to validate data and has ISS equal to or greater than 9, was found to be 65.4%. The development model using the 70% data set, this time with the inclusion of the predictor variables, yielded a prediction accuracy of 71.4%. The validation of the development model was performed on the 30% data set, which produced 73.1% accuracy. The outcome of logistic regression modeling is presented in Table 7.

Table 7 - Logistic Regression Modeling Results

		Naïve Model						Development Model					
		Model Development Set			Validation Set			Model Development Set			Validation Set		
		ISS			ISS			ISS			ISS		
	Observed	<9	≥9	%	<9	≥9	%	<9	≥9	%	<9	≥9	%
ISS	<9	590	0	100.0	270	0	100.0	508	82	86.1	232	38	85.9
	≥9	370	0	0.00	143	0	0.00	193	177	47.8	73	70	49.0
	Overall %			61.5			65.4			71,4			73.1

The final developed model's log likelihood value (1099.112) was found to be smaller than that of the Naïve model (1279.991), which means that the development model is more accurate in predicting the injury severity score (ISS<9 vs. ISS>9) than the Naïve model.

Finally, based on a p value of .817 from the Hoshmer Lemeshow chi-squared test, which is greater than 0.05, the development model shows a good fit with the data.

In the light of these results, it can be stated that the four significant factors (work experience in project, type of injury, unsafe act and unsafe condition) are valid predictor variables that can be used in estimating whether an accident is going to result in a high injury severity or not. The negative signs of the regression (beta) coefficients listed indicate that work experience in project has decreasing effects on the probability of a workplace injury resulting with high ISS. In other words, the more experience the worker has, the less likely he/she will be exposed to a severe workplace related injury. Faulty tool/equipment/machinery and not using Personal Protective Equipment (PPE) categories from unsafe condition and unsafe act variables respectively had the most significant effect on the ISS. Types of injury such as falls, vehicle accidents, heat exposure and electric shock had increasing effects on ISS.

6. SUMMARY AND CONCLUSION

The study presented in this paper intended to use logistic regression analysis to develop a statistical model in identifying factors that contribute to severity of construction injuries and predicting their severity scores. An extensive archival study was initially performed to establish a database to realize this study. Research methodology included using statistical analysis methods such as frequency analysis (univariate) and cross tabulation (bivariate) analyses before undertaking logistic regression.

Although, univariate and bivariate analyses were primarily used as tools to determine statistically significant variables for logistic regression, the following conclusions can be drawn from the findings:

- Cold weather conditions could have a negative impact on the frequency of injuries.
- “Monday effect” theory was once again confirmed by this study.
- “Falls” and “struck by” were identified as dominant injury types in Turkish construction industry.
- As the age of worker increases, it was observed that frequency of work related injuries decreases. This could be an indicator of work experience positively affecting safety awareness.
- Formwork workers tend to have more injuries than other worker trades.
- Unsafe act and condition variables revealed that safety management or lack thereof is a major problem in Turkish construction sites. Particularly insufficient or unsafe work procedures and not using PPE were encountered the most in accident reports.

  The outcome of the logistic regression analysis showed that four of the independent variables (work experience, accident type, unsafe condition and unsafe act) have statistically significant influence on workplace injury severity. To elaborate further on the regression results;

- Workers who have more work experience than others are less likely to be exposed to a severe workplace related injury.

- Faulty tool/equipment/machinery and not using Personal Protective Equipment (PPE) categories from unsafe condition and unsafe act variables respectively had the most significant effect on the severity of construction injuries. In other words, injuries caused by those unsafe acts and conditions are more likely to end up severe.
- Types of injury such as falls, vehicle accidents, heat exposure and electric shock have higher probability of causing severe injuries.

7. DISCUSSION AND RECOMMENDATIONS

Injury severity score (ISS), which had not been applied in the construction safety related research studies before, was identified as the dependent variable in the logistic regression analysis part of the study. Using ISS values as a dependent variable and including scoring in the model has been one of the features that made this study unique. In addition, the nonlinear nature of the ISS allowed the use of this scoring system as a dependent variable in the logistic regression analysis, which has a similar nonlinear structure. Using ISS makes it possible to compare and rank injuries by their severity and could be used in future studies that involve estimating the costs of injuries and risk analysis.

Research information on modeling the accident outcomes is scarce in safety and health literature. In fact, this was one of the main driving forces of this study that led to selection of binary logistics regression analysis, which is usually used in research areas other than construction safety. Logistic regression analysis developed a model that consists of four multilevel independent variables (work experience in project, type of injury, unsafe condition, and unsafe act), all of which have significant relationships with the binary dependent variable (injury severity score). The improvement of accuracy from the naïve to the development model (9.9%), and further improvement of accuracy on the validation model (7.7%) are indicative of the successful validation of the logistic regression model presented.

The model developed in the study allows to predict the injury severity (whether the ISS is greater or less than 9) by using the information coming from independent variables included in the model and to approach proactively. It would be possible to at least reduce the severity of potential injuries by focusing on the factors that meaningfully contribute to the outcome of injury; such as, increasing training efforts for workers with less work experience, eliminating fall and struck by falling object hazards, providing PPE and encouraging their usage, enforcing written safety procedures.

This study was realized by investigating more than two thousand construction accident case reports that occurred in three big cities (İstanbul, Ankara and İzmir) of Turkey over a span of three years with the special permission of Social Security Institution. It is expected that the database created for this study could be useful in other statistical studies involving Turkey, where accurate statistical injury data is still difficult to obtain. Furthermore, a similar study could be carried out in different locations or with different sectors and comparisons could be made with the findings of this study.

References

[1] Aksorn, T., Hadikusumo, B.H.W. Critical success factors influencing safety program performance in Thai construction projects. Safety Science, 46, 709-727, 2008.

[2] Zeng, S.X., Vivian, W.Y., Tam, C.M. Towards occupational health and safety systems in the construction industry of China. Safety Science, 46, 1155-1168, 2008.

[3] Gillen, M., Baltz, D., Gassel, M., Kirsch, L., Vaccaro, D. Perceived safety climate, job demands, and coworker support among union and nonunion injured construction workers. Journal of Safety Research 33, 33-51, 2002.

[4] Fass, S., Yousef, R., Liginlal, D., Vyas, P., Understanding causes of fall and struck-by incidents: What differentiates construction safety in the Arabian Gulf region?, Applied Ergonomics, 58, 515-526, 2017.

[5] Pérez-Alonso, J., Carreno-Ortega, A., Vázquez-Cabrera, F.J., Callejón-Ferre, A.J. Accidents in the greenhouse-construction industry of SE Spain. Applied Ergonomics, 43, 69-80, 2012.

[6] Unsar, S., Sut, N. General assessment of the occupational accidents that occurred in Turkey between the years 2000 and 2005. Safety Science, 47, 614-619, 2009.

[7] Haslam, R.A., Hide, S.A., Gibb, A.G.F., Gyi, D.E., Pavitt, T., Atkinson, S., Duff, A.R. Contributing factors in construction accidents. Applied Ergonomics, 36(4), 401-415, 2005.

[8] Çolak, B., Etiler, N., Biçer, U. Fatal Occupational Injuries in the Construction Sector in Kocaeli, Turkey, 1990-2001. Industrial Health, 24, 424-430, 2004.

[9] Spangenberg, S., Baarts, C., Dyreborg, J., Jensen, L., Kines, P., Mikkelsen, K.L. Factors contributing to the differences in work related injury rates between Danish and Swedish construction workers. Safety Science 41, 517-530, 2003.

[10] Tözer, K.D., Çelik, T., Gürcanlı, G.E., Classification of Construction Accidents in Northern Cyprus. Teknik Dergi, DOI: 10.18400/tekderg.325546, 29 (2), 8295-8316, 2018.

[11] Bilir, S., Gürcanli, G.E., A Method For Determination of Accident Probability in Construction Industry. Teknik Dergi, DOI: 10.18400/tekderg.363613, 29 (4), 8537-8561, 2018.

[12] Baradan, S., Akboga, O., Cetinkaya, U., Usmen, M., Ege Bölgesindeki İnşaat İş Kazalarının Sıklık ve Çapraz Tablolama Analizleri. Teknik Dergi, 27 (1), 7345-7370, 2016.

[13] Al-Ghamdi, A.S. Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. Accident Analysis and Prevention, 34, 729-741, 2002.

[14] Mercier, C.R., Shelley, II M.C., Rimkus, J.B., Mercier, J.M. Age and Gender as Predictors of Injury Severity in Head-on Highway Vehicular Collisions. Transportation Research Record: Journal of the Transportation Research Board, 1581, DOI: 10.3141/1581-05, 1997.

[15] Kim, K., Li, L., Modeling Fault among Bicyclists and Drivers Involved in Collisions in Hawaii, 1986–1991. Transportation Research Record: Journal of the Transportation Research Board, 1538, 75-80, 1996.

[16] Steyeberg, E.W., Harrell, Jr F.E., Borsboom, G.J., Eijkemans, M.J., Vergouwe, Y., Habbema, J.D. Internal Validation of Predictive Models: Efficiency of Some Procedures for Logistic Regression Analysis. Journal of Clinical Epidemiology, 54: 774-781, 2001.

[17] Li, Y., Bai, Y. Development of Crash-Severity-Index Models for the Measurement of Work Zone Risk Levels. Accident Analysis and Prevention, 40, 1724-1731, 2008.

[18] Fang, D., Chen, Y., Wong, L. Safety Climate in Construction Industry: A Case Study in Hong Kong. Journal of Construction Engineering and Management, 132, 573-584 2006.

[19] Cakan, H., Kazan, E., Usmen, M. Investigation of Factors Contributing to Fatal and Nonfatal Roofer Fall Accidents. International Journal of Construction Education and Research, 10, 300-317, 2014.

[20] Irizarry, J., Abraham, D.M. Assessment of Risk Perception of Ironworkers. Journal of Construction Research, 7(1/2), 111-132, 2006.

[21] Baradan, S., Usmen, M. Comparative Injury and Fatality Risk Analysis of Building Trades. Journal of Construction Engineering and Management, 132 (5), 533-539, 2006.

[22] Social Security Institution (SSI) Official Statistics http://www.sgk.gov.tr/wps/portal/sgk/tr/kurumsal/istatistik/sgk_istatistik_yilliklari Accessed February 07, 2018.

[23] Hürriyet, Daily News, 10 workers killed in construction site accident in Istanbul, http://www.hurriyetdailynews.com/10-workers-killed-in-construction-site-accident-in-istanbul-.aspx?pageID=238&nID=71381&NewsCatID=341 Accessed February 02, 2018.

[24] The Telegraph, At least 11 die in fire at construction site in Turkey, http://www.telegraph.co.uk/news/worldnews/europe/turkey/9137516/At-least-11-die-in-fire-at-construction-site-in-Turkey.html Accessed February 01, 2018.

[25] Act No 6331, Occupational Health and Safety Law, Date of Enactment: 20/06/2012.

[26] Tabachnick, B.G., Fidell, L.S. Using Multivariate Statistics. 6th Ed., Pearson Education. Inc., Boston, MA. 2013.

[27] Kass, G.V. An exploratory technique for investigating large quantities of categorical data. Applied Statistics, 29:2, 119–127, 1980.

[28] Suruda, A., Fosbroke, D., Braddee, R. Fatal work-related falls from roofs. Journal of Safety Researc, 26(1), 1-8, 1995.

[29] Stern, F.B., Ruder, A.M., Chen, G. Proportionate mortality among unionized roofers and waterproofers. American Journal of Industrial Medicine, 37(5), 478-492, 2000.

[30] Huang, X., Hinze, J. Analysis of Construction Worker Fall Accidents. Journal of Construction Engineering and Management, 123(3), 262-271, 2003.

[31] Beavers, J.E., Moore, J.R., Schriver, W.R. Steel Erection Fatalities in the Construction Industry. Journal of Construction Engineering and Management, 227-234, 2009.

[32] Im, H.J., Kwon, Y.J., Kim, S.G., Kim, Y.K., Ju, Y.S., Lee, H.P. The characteristics of fatal occupational injuries in Korea's Construction Industry. Safety Science, 47, 1159-1162, 2009.

[33] Gürcanlı, E., Müngen, U. Analysis of construction accidents in Turkey and responsible parties. Industrial Health, 51, 581-598, 2013.

[34] ICD, International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010, Chapter XX External causes of morbidity and mortality (V01-Y98), W00-X59 Other external causes of accidental injury, http://apps.who.int/classifications/icd10/browse/2010/en#/XX Accessed May 03, 2018.

[35] ILO, 2014, "Theory of Accident Causes", http://www.iloencyclopaedia.org/part-viii-12633/accident-prevention/92-56-accident-prevention/theory-of-accident-causes Accessed April 21, 2018.

[36] Hill, D.C. Construction Safety Management and Engineering. American Society of Safety Engineers, Des Plaines, Illinois, 707 p. 2004.

[37] DMI, Disaster Management Institute, "The Domino Theory", http://www.hrdp-idrm.in/e5783/e17327/e24075/e27357/ Accessed January 24, 2017.

[38] Linn, S. The Injury Severity Score-Importance and Use. Ann Epidemiol., 5(6), 440-446, 1995.

[39] Baker, S.P. Evaluation of Medical Care of the Injured. Journal of Trauma, 11, 892-894, 1971.

[40] FTR, Florida Trauma Registry, A Report on The Collection Of Trauma Patient-Level Data From Trauma Centers And Other Acute Care Facilities For 2002. 2002 Annual Report, Division of Emergency Medical Operations Florida Department of Health July, 2003.

[41] Baker, S.P., O'Neill, B. The Injury Severity Score: An Update. The Journal of Trauma, 16(11), 882-885, 1976.

[42] Krischer, J.P. Indexes of Severity: Underlying concepts. Health Serv. Res., 11, 143-157, 1976.

[43] Krischer, J.P. Indexes of Severity: "Conceptual Development. Health Serv. Res., 14, 56-57, 1979.

[44] Sacco, W.J., Jameson, J.W., Copes, W.S., Lawnick, M.M., Keast, S.L., Champion, H.R. Progress toward a New Injury Severity Characterization: Severity Profiles. Comput. Biol. Med., 18, 419-428, 1988.

[45] Copes, W.S., Champion, H.R., Sacco, W.J., et al., Progress in Characterizing Anatomic Injury. Journal of Trauma, 30, 1200-1207, 1990.

[46] Rainer, T.H., Chan, S.Y., Suen, D.T.K., Lam, W., Cocks, R.A. Severe Trauma Presenting to the Resuscitation Room of a Hong Kong Emergency Department. Hong Kong Journal of Emergency Medicine, 7(3), 129-135, 2000.

[47] Ballesteros, M.F., Dischinger, P.C., Langenberg, P. Pedestrian Injuries and Vehicle Type in Maryland, 1995-1999. Accident Analysis and Prevention, 36, 73-81, 2004.

[48] Moore, S.J. The Effects of Daily Snowfall Accumulation on the Pattern and Severity of Traumatic Injuries from a U.S. Ski Resort. Trauma/Critical care Services, Vail Valley Medical Center, Colorado, 2011.

[49] Moylan, J.A., Detmer, D.E., Rose, J., et al:, Evaluation of the Quality of Hospital Care for Major Trauma. Journal of Trauma, 16, 517-523, 1976.

[50] Gillen, M., Faucett, J.A., Beaumont, J.J., McLoughlin, E. Injury Severity Associated with Nonfatal Construction Falls. American journal of Industrial Medicine, 32, 647-655, 1997.

[51] Glancy, K.E., Glancy, C.J., Lucke, J.F., et Al, A Study of Recovery in Trauma Patients. F., Trauma, 33, 602-609, 1992.

[52] Richardson, D., Loomis, D., Bena, J., Bailer, A.J. Fatal occupational injury rates in southern and non-southern states, by race and Hispanic ethnicity. Am. J. Public Health 94, 1756–1761, 2004.

[53] Ringen, K., Englund, A., Welch, L., Weeks, J.L., Seegal, J.L. Why construction is different. Occupational Medicine: State of the Art Reviews, 10(2), 255– 259, 1995.

[54] Ringen, K., Seegal, J.L., Englund, A. Safety and health in the construction industry. American Journal of Public Health, 16, 165–188, 1995.

[55] Babbie, E.R. The Basics of Social Research. 5th Ed., Wadsworth Publishing, California, 2011.

[56] Greenland, S. Modeling and variable selection in epidemiologic analysis. American Journal of Public Health, 79(3), 340- 349, 1989.

[57] Rehm, J., Greenfield, T.K., Walsh, G., Xie, X., Robson, L., Single, E. Assessment methods for alcohol consumption, prevalence of high risk drinking and harm: a sensitivity analysis. International Journal of Epidemiology, 28(2), 219-224, 1999.

[58] Arias, J.C., Miller, R. Market analysis of student' attitudes about credit cards. Business Intelligence Journal, 3(1), 23-36, 2010.

[59] Rollins, B.C., Feldman, H. Marital Satisfaction over the Family Life Cycle. Journal of Marriage and Family, 32(1), 20-28, 1970.

[60] Sureshchandar, G.S., Rajendran, C., Anantharaman, R. The Relationship between Service Quality and Customer Satisfaction–A Factor Specific Approach. Journal of Services Marketing, 16(4), 363-379, 2002.

[61] Fields, A. Discovering Statistics Using SPSS. 2nd Ed., Sage Publications, Thousand Oaks, CA, 2005.

[62] Sims, R.L. Bivariate Data Analysis: A Practical Guide. Nova Science Pub., Inc. 1999.

[63] Elliott, A.C., Woodward, W.A. Statistical Analysis Quick Reference Guidebook with SPSS Examples. 1st Ed., SAGE Publications, Inc., California. 2007.

[64] Healy, J. Statistics: a tool for social research. Wadsworth Publishing, 9th edition. 2011.

[65] Hosmer, W.D., Lemeshow, S. Applied logistic regression. John Wiley & Sons, Hoboken, NJ. 2000.

[66] Kleinbaum, D.G. Logistic regression: A self-learning text. Springer-Verlag; 1st edition, New York, NY. 1994.

[67] Kazan, E.E. Analysis of Fatal and Nonfatal Accidents Involving Earthmoving Equipment Operators and On-Foot Workers. PhD Thesis, Wayne State Universty, 173p. 2013.

[68] Arboretti Giancristofaro, R., Salmaso, L. Model Performance Analysis and Model Validation in Logistic Regression. Statistica LXIII, 375 – 396, 2003.

[69] Azen, R., Walker, C.M. Categorical Data Analysis for the Behavioral and Social Sciences. 1st Ed., Routledge Academic, New York. 2010.

[70] Campolieti, M., Hyatt, D.E. Further Evidence on the "Monday Effect" in Workers' Compensation. Industrial and Labor Relations Review", 59(3), 438-450, 2006.

[71] Chi, C.F., Chang, T.C., Ting, H.I. Accident patterns and prevention measures for fatal occupational falls in the construction industry. Applied Ergonomics, 36 (4); 391-400, 2005.

[72] Kemmlert, K., Lundholm, L. Slips, trips and falls in different work groups—with reference to age and from a preventive perspective. Applied Ergonomics, 32(2), 149-153, 2001.

[73] Ore, T., Stout, N. Traumatic occupational fatalities in the US and Australian construction industries. Am J Ind Med 30, 202–206, 1996.

[74] Hinze, J., Pedersen, C., Fredley, J. Identifying root causes of construction injuries. Journal of Construction Engineering and Management, 124(1), 67-71, 1998.

[75] Jackson, S.A., Loomis, D. Fatal Occupational Injuries in the North Carolina Construction Industry, 1978–1994. Applied Occup Environ Hyg 17, 27–33, 2002.

[76] Fabrega, V., Stakey, S. Fatal Occupational Injuries among Hispanic Construction Workers of Texas, 1997 to 1999. Human and Ecological Risk Assessment, 7, 1869–1883, 2001.

[77] Arndt, V., Rothenbacher, D., Daniel, U., Zschenderlein, B., Schuberth, S., Brenner, H. All-cause and cause specific mortality in a cohort of 20,000 construction workers; results from a 10 year follow up. Occup. Environ. Med. 61(5), 419-425, 2004.