



## Ranking and Determining the Factors Affecting the Road Freight Accidents Model

Masoud Bagheri Ramiani <sup>a\*</sup>, Gholamreza Shirazian <sup>b</sup>

<sup>a</sup> PhD Student of Transportation, Department of Civil Engineering, Shomal University, Amol, Iran.

<sup>b</sup> PhD of Transportation Engineering, Department of Civil Engineering, Shomal University, Amol, Iran.

Received 15 November 2019; Accepted 04 March 2020

### Abstract

The tremendous growth of population, particularly in developing countries, has led to increased number of travels, especially those with load and freight specifications. Hence, expanding the present facilities or developing new networks or systems concerning freight and transportation is an essential issue. Among the various transportation systems, road freight has secured a significant place in sub-urban transportation, as it is responsible for transporting loads, decreasing transportation costs, and increasing the safety of highway users. Besides these advantages, poor and nonstandard design and performance of sub-urban highways and transport fleet and equipment leads to the increased number of accidents and inefficiency of these facilities. Based on these facts, the primary aim of the present study is to probe into the factors affecting road freight accident severity. For this purpose, the data obtained from road freight accidents occurring in 2016, 2017, and 2018 in Gilan Province, Iran, were used for analyzing the frequency, ranking and determining the factors, and creating models for accident severity. The results indicated that in accordance with the accident severity model in 2016, several factors such as the season of autumn, daytime light, drivers aged from 18 to 60, and pickup trucks have impacted the on-road freight accident severity. While, in 2017 the severity was affected by factors like rural road, freight trucks, non-faulty passenger cars, motorcycles, and pedestrians. When considering the effective variables in 2018, it was found that such factors as the accident time (usually occurring between 12 p.m. to 6 p.m.), rural and major roads, freight trucks, non-faulty motorcycles, and the careless driving without due care and attention to the front were the variables affecting road freight accidents. Moreover, not following safety guidelines during freighting is the most effective variable in road freight accidents.

**Keywords:** Accidents; Road Freight; Cargo; Damage; Injury; Fatality.

### 1. Introduction

Over the last decade and also nowadays, road accidents have experienced an increasing trend; and there have been various studies on accidents around the world [1-3]. Road accidents are considered to be a typical phenomenon all over the world and approximately 1.3 million civilians die as a result of this phenomenon. Moreover, approximately 20 to 50 million people have been injured in these accidents, where the majority were young people with ages ranging from 15 to 45. Road accidents are claimed to be the ninth most important factor of fatality in the world as it accounts for 2.2% of the mortality rate in the world. The costs of accidents are estimated to be about 500 million dollars all over the world. This is equal to between 1 to 2 percent of GDP in the countries with low to average income. The current trend of accidents indicates that if emergent measures are not taken in this regard, it is likely that road injuries will

\* Corresponding author: [mbrpost@yahoo.com](mailto:mbrpost@yahoo.com)

<http://dx.doi.org/10.28991/cej-2020-03091518>



© 2020 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

become the seventh most important factor of fatality by 2030, 90% of which will have occurred in the countries with low to average income. It is essential that emotional and mental injuries as well as permanent disabilities resulting from these accidents be added to these complications [4].

Analyzing the accidents, it could be concluded that normally several parameters account for the accidents, which can be categorized into human, environmental, and vehicle-dependent factors [5]. Various studies have revealed that some parameters like AADT, traffic congestion, exceeding the speed limit and the number of lanes, drivers' distraction, and weather conditions are among the factors which affect accidents [6-8]. Increase in the severity of any of the above-mentioned factors and/ or any inappropriateness of these conditions can lead to increased number of accidents [9-11].

An increase in the number of accidents results in the increase in the number of owners of motor vehicles. It can be said that the mentioned growth has been about 65% over the last 2 decades; while in developing countries it has occurred faster than this [12]. Because of this, road accidents are considered to be one of the most significant issues of public health in any given society. Besides, it can be concluded that this problem is more serious than other public health-related issues as the majority of its victims are young and healthy people [13, 14].

Ghaffar et al. (2004) [15] evaluated the effects of road traffic injuries (RTI) in Pakistan. The results indicated that most accidents happened between 12 and 18 pm. The level of RTI is higher among people aged between 16 and 45 years. Furthermore, the results showed that RTI is almost three times higher in males than females.

Labinjo et al. (2009) [16] provided a population-based survey to explore the epidemiology of RTI in Nigeria. The results showed that motorcycle accidents accounted for 54.33% of all RTI. The risk of crashes was higher among males aged between 18 and 44.

Hu et al. (2012) [17] explored the characteristics of traffic accidents on rural roads using the quantitative analysis. The research shows that 92.68% and 5.42% of casualties occurred on tangents and curves, respectively. Considering the time parameter, casualties during the daytime have been more serious than the night. Also, the results showed that crashes that cause injuries are most common during the day rather than night and motor vehicle accidents account for the majority of casualties. Zangoeei Dovom et al. (2013) [18] explored the fatal accident distribution in Mashhad, Iran. According to the results, the male had more fatalities than the females. For both genders, most accidents had a peak at the ages of 21-30. The male to female overall casualty ratio was 3.41. Among all the road users, the riskiest group was male motorcyclists.

Lee and Jeong (2016) [19] investigated the characteristics of traffic collisions occurred in expressways and rural roads among the truck drivers. The results showed that with respect to the day of the week, the accident rate was higher in the middle of the week. On rural roads, the accident rate during the daytime was much higher than the night time (81.7%). The accidents occurred mostly in clear/cloudy weather (76.2%). Besides, the majority of accidents occurred over a straight road (62.2%), followed by an intersection (15.4%) and a curved road (9.4%).

Road freight is the oldest way of transporting cargo in the world. In terms of the price and speed, it is the most appropriate way to transport a variety of goods and cargo. Nowadays, the majority of the cargo is transported by means of road freight, which along with its merits, has some disadvantages. One of the disadvantages is the accidents occurring due to the failure of freight vehicles on roads, which in addition to economic losses results in fatality or injury of road users as well. Hence, in this study, the effective factors are identified by means of considering the severity of road freight accidents, particularly the accidents resulting from some vehicles like pickup trucks, trucks, trailer cars, etc. Then, through modelling and statistical analyses of accidental data, the appropriate approaches to decrease road accidents and increase civil welfare and traveller's safety are identified. For this purpose, after analyzing the frequency of accidents, their ranking is performed and the factors affecting the severity of accidents are determined. Moreover, the impact of independent variables on the severity of freight vehicle accidents is modelled. The purposes and adopted innovations in the present study can be summarized as follows:

- Application of Friedman test and Factor analysis methods for the road freight accidents,
- Investigating the effect of independent variables on the severity of freight vehicles,
- Frequency analysis of variables affecting freight vehicles accidents,
- Ranking independent variables affecting the severity of freight vehicles,
- Modelling of independent variables affecting the dependent variables of freight vehicles accidents severity,

## 2. Research Methodology

In this section, initially, the specifications of the study area will be introduced. Then, different statistical analysis methods and modelling of accident severity will be executed according to Figure 1.

## 2.1. Case study

Gilan is one of the northern provinces of Iran, whose capital is Rasht megacity. This province lies along the Caspian Sea and Azerbaijan, sharing with it an international boarder via Astara in the north. It is located to the west of Mazandaran Province, east of Ardabil Province and north of Zanjan and Qazvin Provinces. Gilan covers an area of 14044 square kilometres and based on the census carried out in 2012, its population is 2480874. Gilan is the tenth province in Iran in terms of population and is the second most populated province in northern Iran i.e. it ranks only second to Mazandaran Province. The population density of this province is 177 persons per square kilometre, which secures third place in Iran. Constituting 46% of the total population of the province, Rasht megacity is the center of the province and the most populated city in the north of the country and the 11<sup>th</sup> most populated city in Iran [20, 21].

Sub-urban highways of Gilan Province are 2573 kilometres in length, of which 1682 km, i.e. 65% of the total length of highways within the province and 2.2% of the total length of highways of Iran, are capable to be utilized for freight purposes. There are 363, 256, and 1063 kilometres of the mentioned total highway network length function as highway, main, and rural roads, respectively. Figure 2 displays all the existing roads in Gilan Province, which can be used for freight transport. Due to the abundance of details, the functions of rural roads are excluded [22].

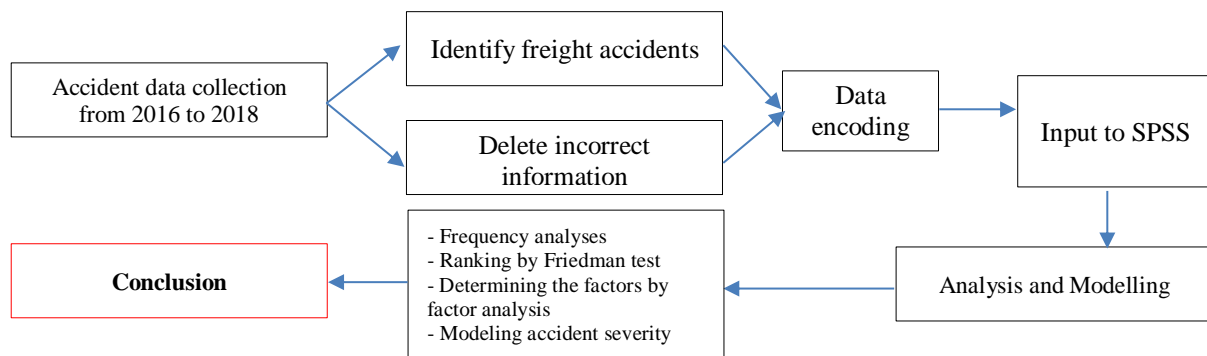


Figure 1. Flowchart of the present study

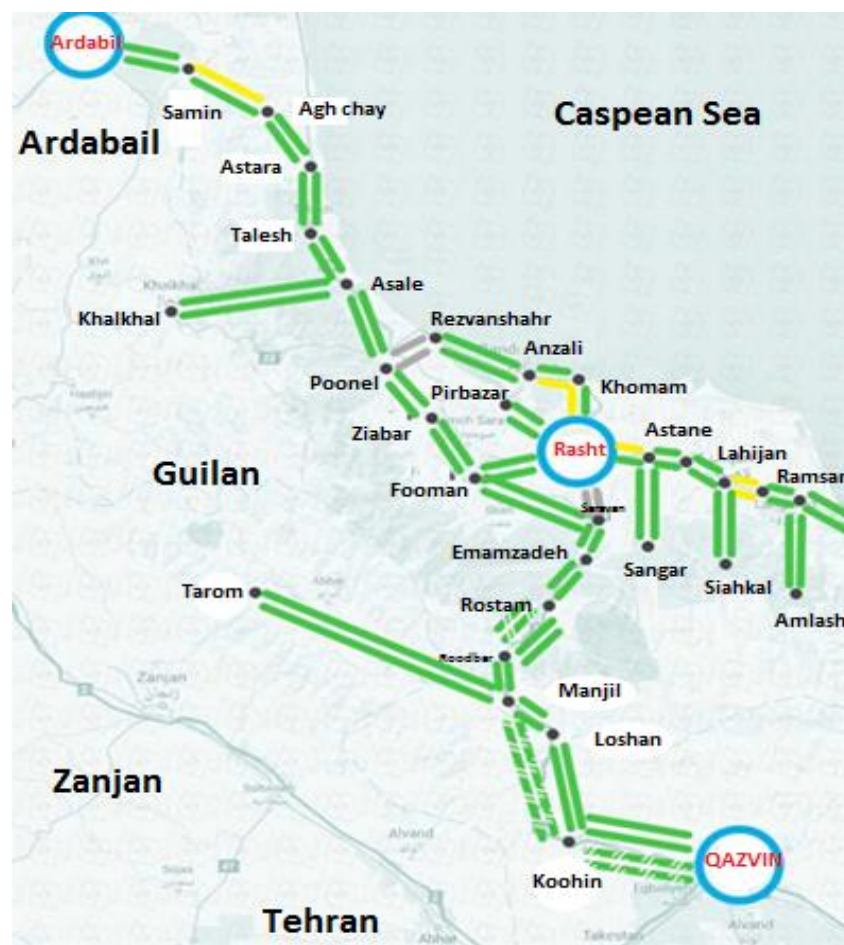


Figure 2. Representing the roads capable of freighting in Gilan Province (Iran)

## 2.2. Statistical Analysis and Modeling Methods

For analysing the Road Freight Accidents of Gilan Province, some statistical analysis and modelling methods, including Kolmogorov–Smirnov test (K–S test), Friedman analysis, Factor analysis, and Logit modelling were used. The statistical analyses used in this study were performed using SPSS software.

### 2.2.1. Kolmogorov-Smirnov Test (K–S test)

One of the main assumptions for most statistical tests is the normality of data distribution where the Kolmogorov–Smirnov test (K–S test) is utilized for this purpose. This test is a nonparametric test for data distribution. In approximate significance test, comparing the output with  $\alpha$  (significance level), the normality of data distribution can be determined. If  $\alpha = 0.05$  (means with 95% certainty) if P-value  $> 0.05$ , the data distribution can be assumed as normal. Indeed this test is a compliance testing of the quantitative data distribution. Normality distribution test is the most common test for examining the normality of a specific distribution [23].

### 2.2.2. Friedman Test

The Friedman test is one of the statistical tests used to compare between several groups and, ranks groups by using the average value, whether these groups belong to one community or not. This test is a non-parametric one corresponding to the F test and is usually used in ranking scales rather than the F test [24]. In the F test, there should be homogeneity of variances that is less observed in ranking scales. The Friedman test is applied for the analysis of two-way variance (for non-parametric data) by a ranking method. Also it is used to compare the average ranking of different groups.

### 2.2.3. Factor Analysis

The factor analysis method is used to find out the underlying variables of a phenomenon or for summarizing a set of data. The primary data for factor analysis is the matrix of correlation between variables. Factor analysis does not have predetermined dependent variables. Factor analysis is applied for two general categories: exploratory purposes and confirmatory purposes. If there is no speculation about the structure of the dimensions relationships, exploratory factor analysis is used. Otherwise, the confirmatory factor analysis is used [25].

In the exploratory factor analysis, the researcher seeks to investigate the empirical data to discover and identify the indices as well as the relationships between them. There is no pre-defined model here. In other words, exploratory analysis, in addition to its exploratory or suggested value can be a structure maker, modeller, or hypothesis creator. Exploratory factor analysis is used when the researcher does not have sufficient previous and pre-empirical evidence to create a hypothesis about the number of underlying factors and wants to determine the number or nature of the factors justifying the covariance of variables. Therefore, exploratory analysis is more considered as a method of compilation and production of a theory, rather than a method of testing a.

### 2.2.4. Multiple Logit Regression

Establishing a relationship between the set of variables  $x$  and the dependent variable  $Y$ , we would encounter a multivariable problem. In analysing such a problem, various types of mathematical models have been used to consider the complexity of the relationship between these variables. The logit regression method is a mathematical method used to describe the relationship between multiple variables denoted by  $x$  and a two-valued dependent variable. A function that is used in this method is an S-shaped function called the logit function, which can also be applied in multi-valued problems by expansion [26]. As it is known, the logit regression method can be utilized to define the variable  $Y$  as the multi-valued parameter. In the simplest case, we can consider  $P(Y=i)$  as a linear function of  $XI(P_i = x_i \beta)$ , where  $\beta$  is the vector of regression coefficients. This equation considers that the probability  $P_i$  at the left side of the equation should be between zero and one, but the linear vector product  $x_i \beta$  at the right side could include all the real numbers. A simple method for solving this problem is to use the probability transfer function to remove the distance limits and model the transferred function as a linear function of the parameters. This conversion occurs in two steps. First, the  $P_i$  probability changes to the chance of success according to Equation 1:

$$OddS = \frac{p_i}{1 - p_i} \quad (1)$$

In the second step, the logarithm of the above-mentioned equation is taken to obtain the logit or success chance logarithm (Equation 2):

$$Logit(p_i) = \text{Log} \frac{p_i}{1 - p_i} \quad (2)$$

The results are quite similar. The reverse transfer function, also called anti-logic, is applied to calculate the probability in terms of logit (Equation 3):

$$\text{Logit}^{-1}(Z_i) = \frac{e^{Z_i}}{1 + e^{Z_i}} \quad (2)$$

The fact that the value of the logit function varies between zero and one, is the first reason for using this function in the probability problems. The second reason in this regard is the form of this function, so that if we start from negative infinity and move to the right, by increasing  $z$ , the value of  $f(z)$  does not change much and remains in the range of zero until we reach the growth threshold. In this range, the value of the function increases rapidly to approach unity, and at this time, the increase of  $z$  does not have much effect on the increase of the function value. Therefore, the logit is a transfer function that associates the probabilities in the interval  $(0, 1)$  with all the real numbers. The negative logit represents a less than 50% probability, and the positive logit represents a more than 50% probability. Thus, the logit model is a general linear model that has a logit transfer function. In other words, the logit of  $P_i$  probability, instead of the probability, follows the linear model [27].

### 3. Results and Analysis

In this part, accidents data in 2016, 2017, and 2018 were used to identify the variables affecting the accidents leading to damage, injury, and fatality when encountering the freight vehicles. Then, the data were considered in terms of statistics and frequency. K-S test, Factor Analysis, Friedman, and Logit test analyses were employed to consider the variables affecting the severity of accidents.

#### 3.1. Frequency Analysis of the Accidents Results

In this study, there is one dependent variable, i.e. accidents severity, and there are 12 independent variables such as time of the accident, day of the accident, season of the accident, road function type, road pavement condition, accident point geometry, lighting status, type of freight vehicle responsible for the accident, age of the faulty driver, type of the non-faulty vehicle, weather condition, the main cause of the accident, where their frequency analysis is presented in the following part. Figures 3 to 15 present the frequency of each of the variables used in this study from 2016 to 2018, split by year.

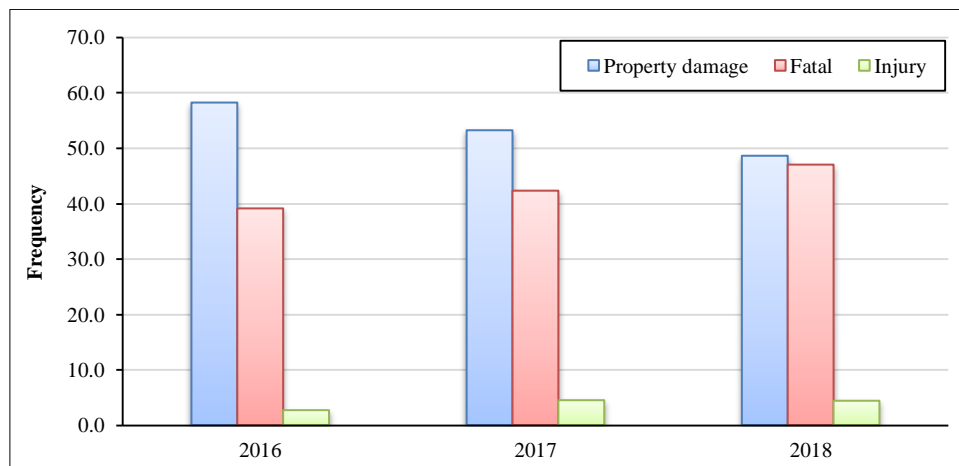


Figure 3. Accident severity frequency (injury, fatality, and damage) by year

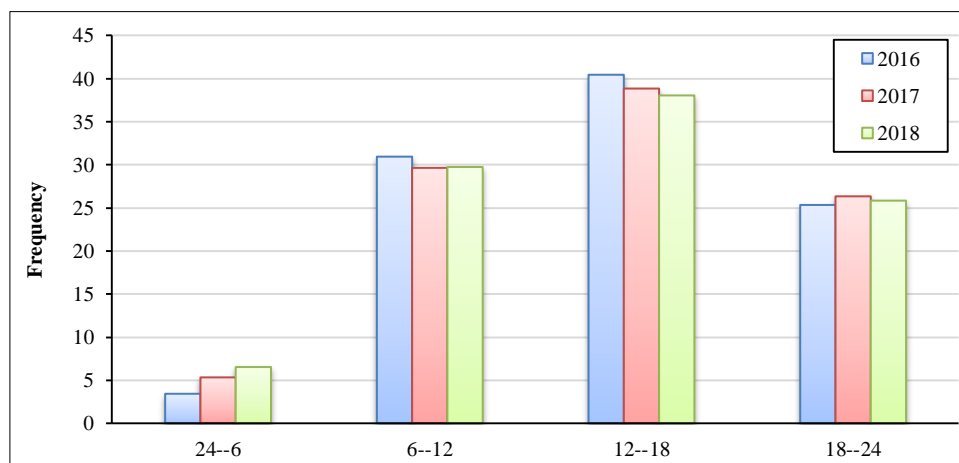


Figure 4. Accidents frequency regarding the time of the accidents by year

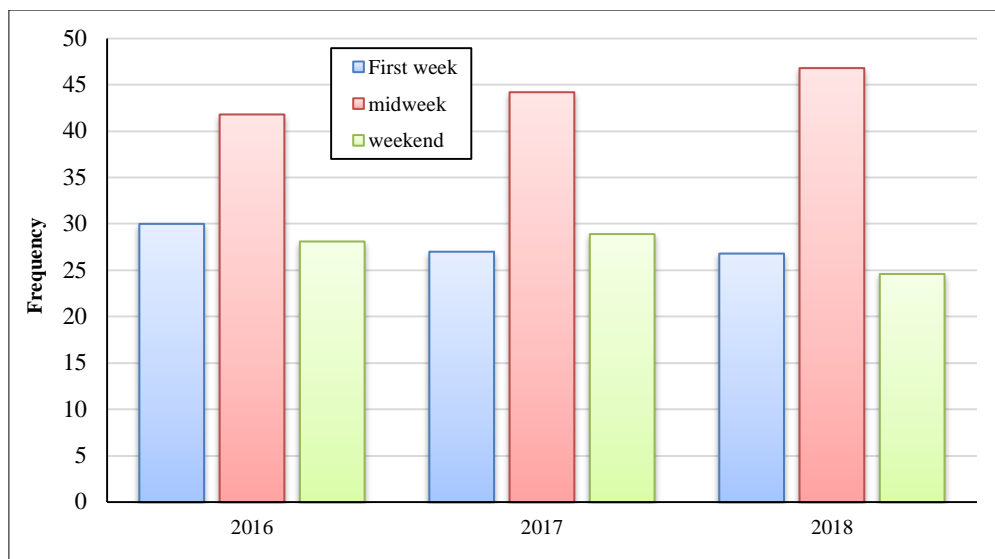


Figure 5. Accidents frequency regarding the day of the accident by year

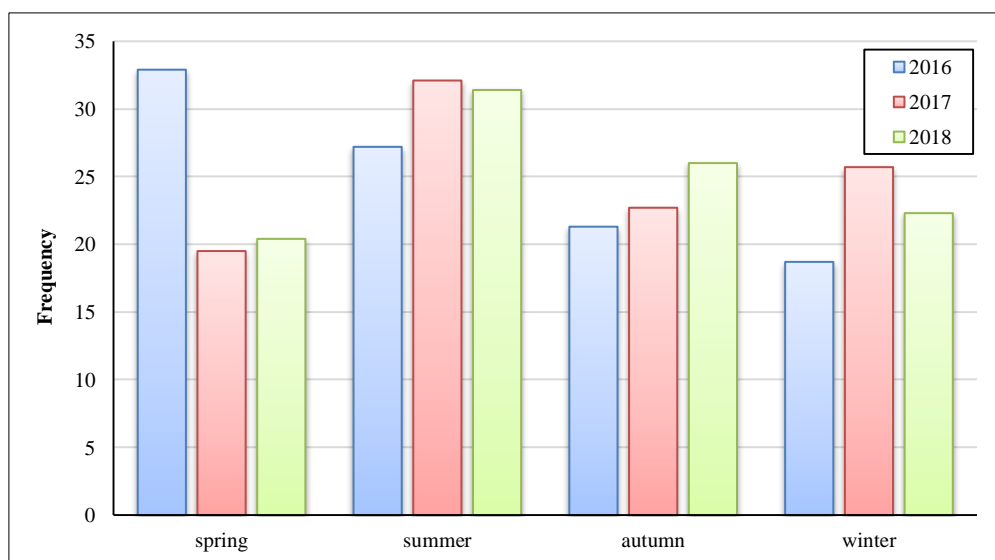


Figure 6. Accidents frequency regarding season of the accident by year

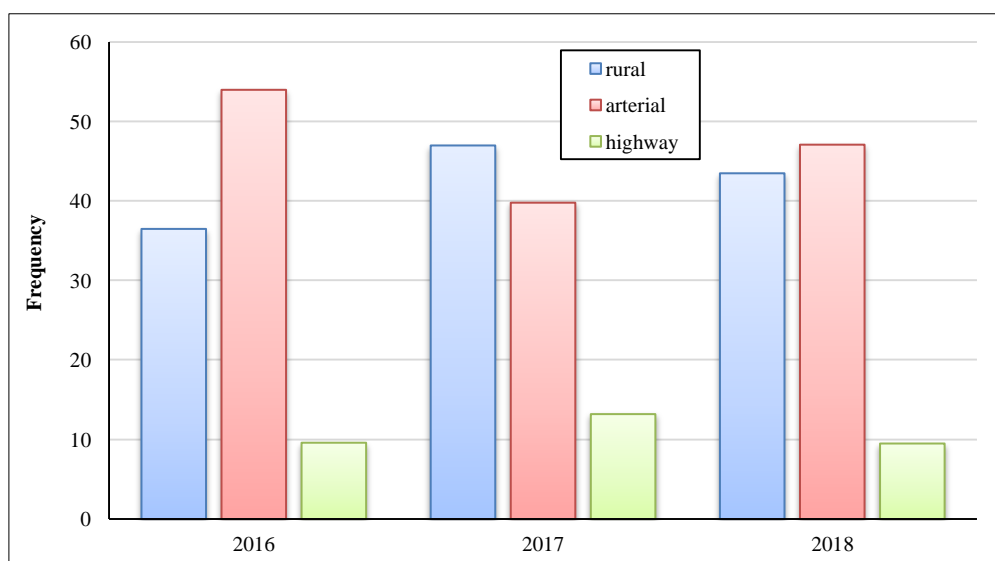


Figure 7. Accident frequency regarding the type of highway function year

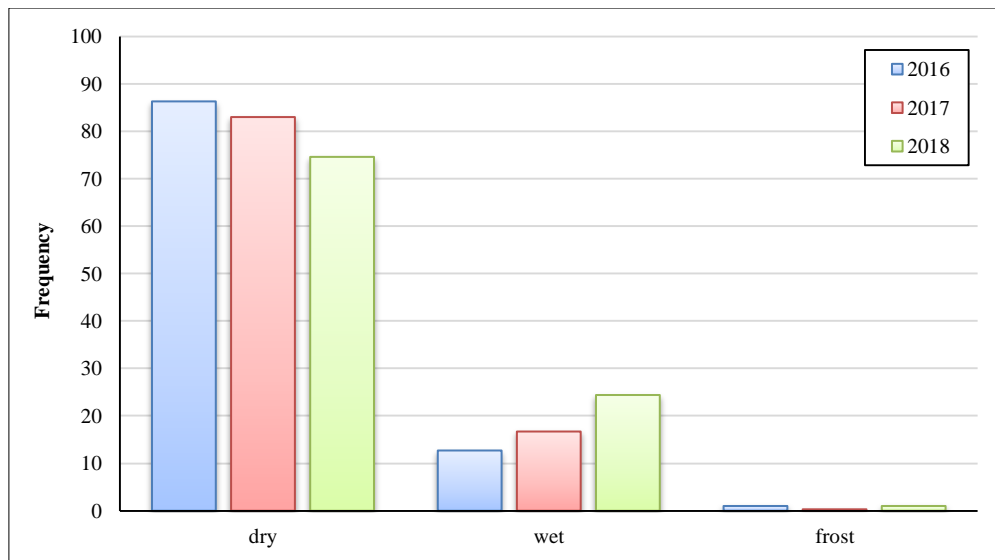


Figure 8. Accidents frequency regarding highway surface condition by year

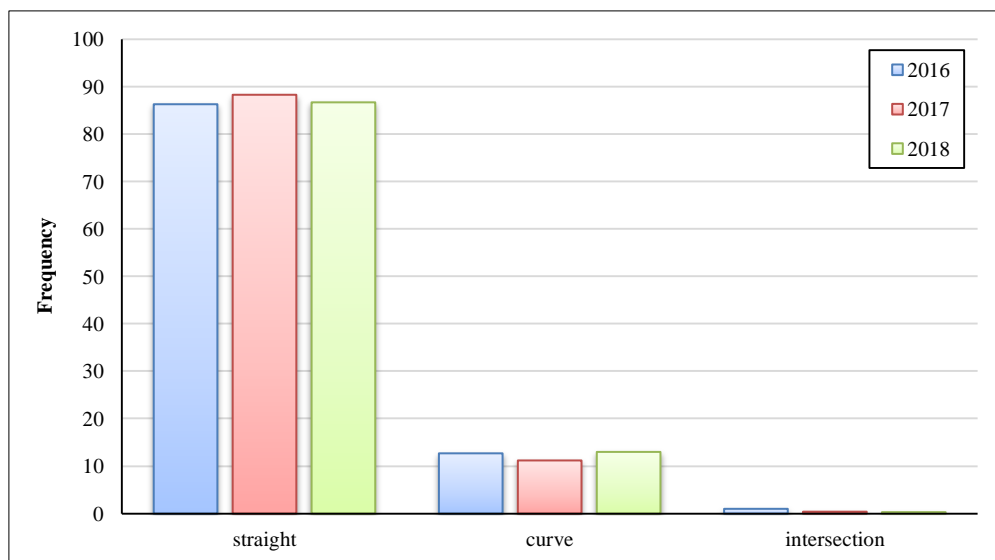


Figure 9. Accidents frequency regarding accident point geometry by year

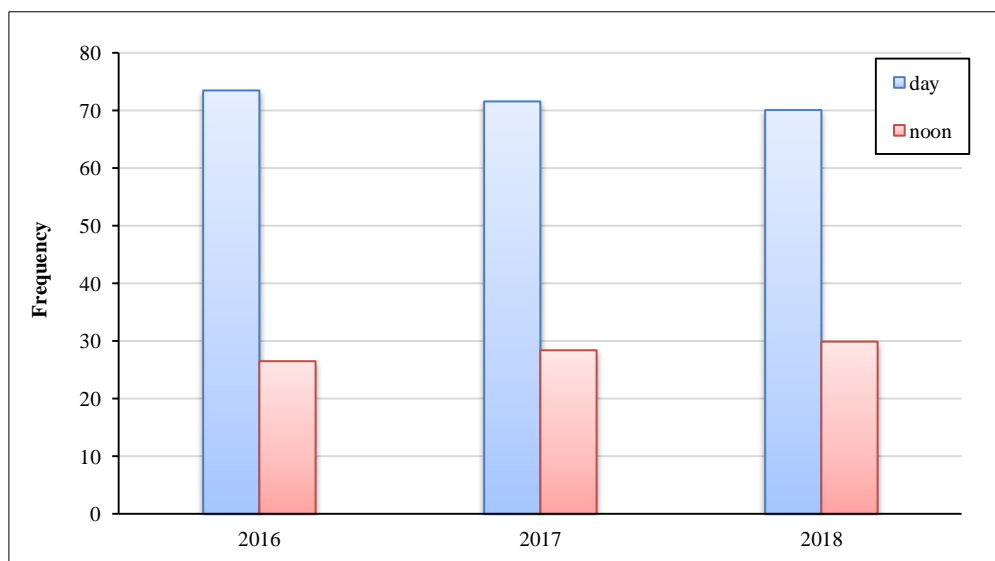
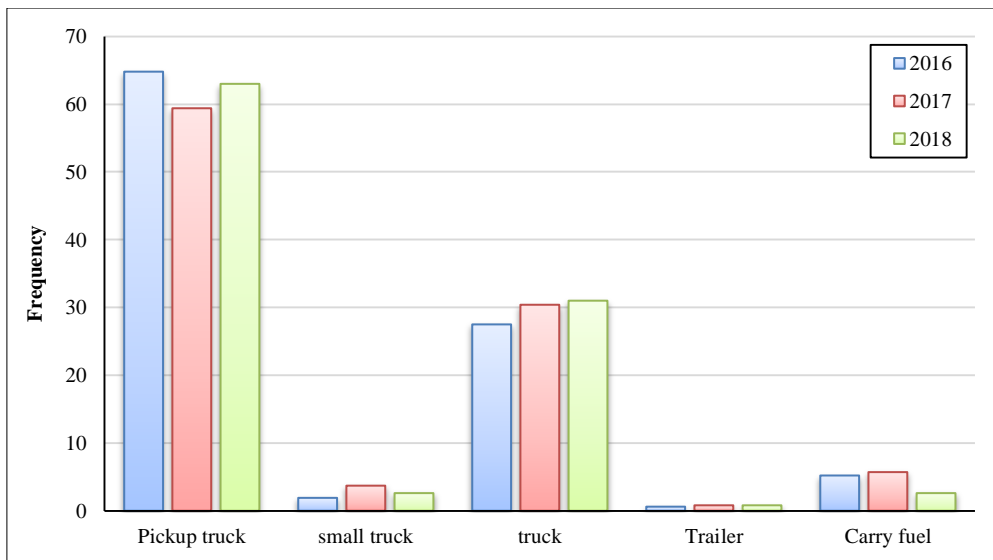
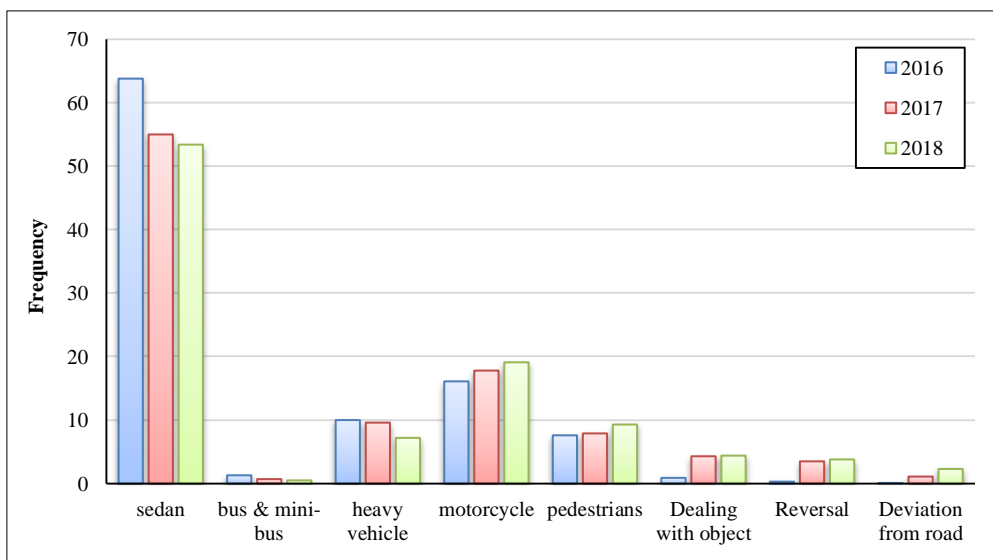


Figure 10. Accidents frequency regarding lighting condition by year

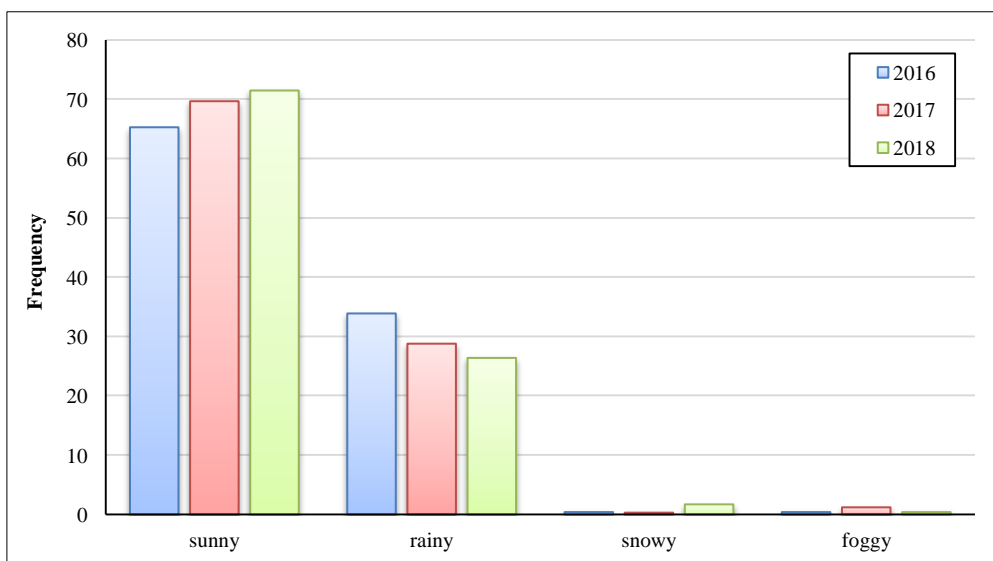




**Figure 11. Accidents frequency regarding the type of the faulty freight vehicle by year**



**Figure 12. Accidents frequency regarding the type of non-faulty vehicle by year**



**Figure 13. Accidents frequency regarding weather condition by year**



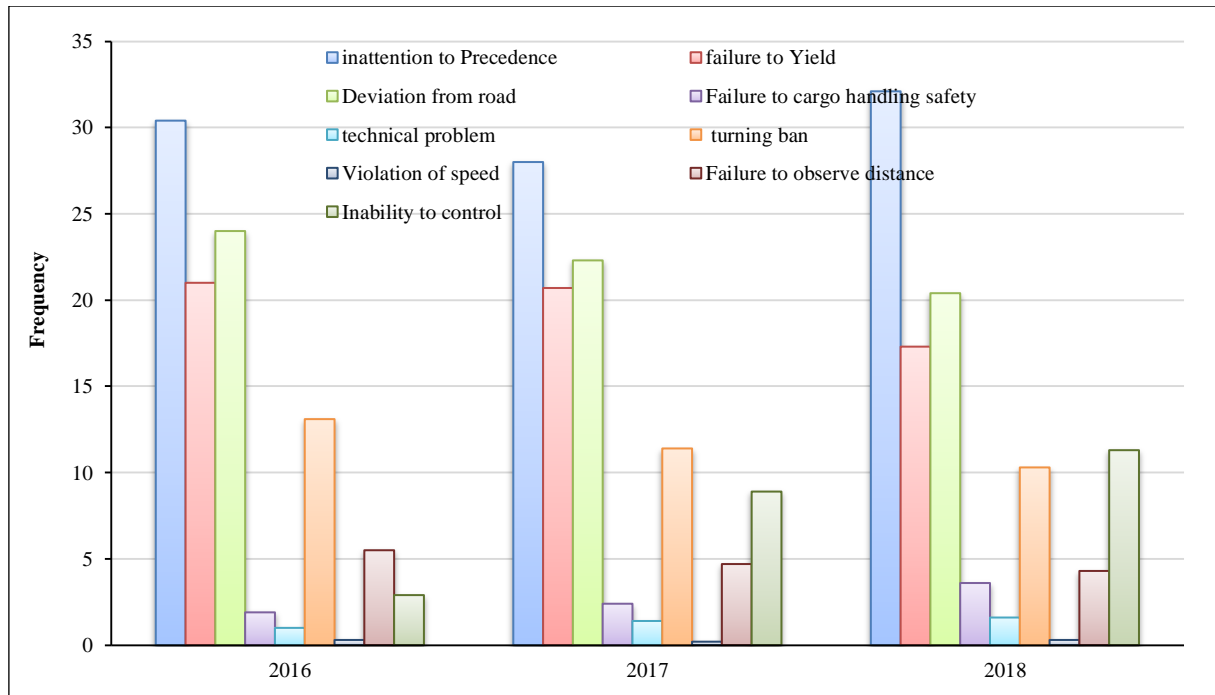


Figure 14. Accidents frequency regarding the type of the main cause by year

The results of Figures 3 to 14 indicate that more than 50% and 40% of road freight accidents are related to the damage and injury accidents, respectively. Meanwhile, the accidents resulting in fatality are only about 10% of the total accidents. Furthermore, the majority of accidents occur between 6 a.m. to 12 a.m and 12 p.m to 6 p.m on weekdays in autumn. In more than 70% of the cases, the accidents occur on daytime light and sunny weather where roads are on a tangent with no curve and good dry pavement condition, which prevent the driver from acceleration. Among all the faulty freight vehicles in the road accidents, pickup trucks and trucks are responsible for 60% and 30% of the accidents, respectively. The other freight vehicles cause 10% of the road freight accidents.

### 3.2. Results of Kolmogorov Smirnov Test

Initially, to select an appropriate test to evaluate the data, it is essential to ensure that there is a normal distribution of statistical data. Thus, K-S test was used to examine whether the distribution is normal. Table 1 summarizes the results of this test. The results indicate that the test is significant for all three years, namely 2016, 2017, and 2018. As a result, these variables do not have a normal distribution and nonparametric tests should be used to make deductions.

Table 1. K-S test results for 2016, 2017, and 2018

Variables	Statistical Test			Significance
	2018	2017	2016	
Time of the accidents	0.214	0.215	0.216	0.00
Season of the accident	0.201	0.211	0.201	0.00
Day of the accident	0.234	0.222	0.210	0.00
Type of highway function	0.444	0.278	0.329	0.00
Highway surface condition	0.267	0.503	0.507	0.00
Accident point geometry	0.331	0.523	0.513	0.00
Lighting condition	0.394	0.452	0.461	0.00
Faulty freight vehicle	0.439	0.367	0.399	0.00
Faulty driver's age	0.289	0.259	0.269	0.00
Non-faulty vehicle	0.518	0.339	0.397	0.00
Weather condition	0.462	0.424	0.411	0.00
Main cause	0.286	0.284	0.284	0.00

### 3.3. Results of Friedman Test

Friedman Test can be utilized to test the equality of variable levels rank. In this study, there are 12 independent variables for ranking the accident severity. Friedman Test is used to determine the rank of each variable. Table 2 presents the data for each variable, Chi-Square statistics range, degrees of freedom, and sig in order. As sig is less than 5%, H0 is rejected and ranking equality (priority) hypothesis of these 12 factors is not accepted. Table 3 also displays descriptive statistics which indicate the mean rank of each variable. The smaller the mean rank, the more important is the corresponding variable.

**Table 2. Results of Friedman parameter**

Year	Number of data	Chi-square	df	Sig.
2016	1152	4565.01	11	0.00
2017	987	3960.32	11	0.00
2018	1189	4702.97	11	0.00

**Table 3. Results of Friedman Test for the accidents**

Independent variables	2016		2017		2018	
	Average	Rank	Average	Rank	Average	Rank
Time of the accidents	9.59	12	9.31	12	9.24	12
Season of the accident	7.65	9	8.34	9	8.21	9
Day of the accident	7.04	7	6.99	7	6.91	6
Type of highway function	8.91	11	8.49	10	8.66	10
Highway surface condition	3.95	2	3.86	2	4.19	2
Accident point geometry	3.88	1	3.63	1	3.66	1
Lighting condition	4.38	3	4.32	3	4.33	3
Faulty freight vehicle	5.81	5	6.02	5	5.69	5
Faulty driver's age	7.35	8	7.15	8	7.05	8
Non-faulty vehicle	6.15	6	6.83	6	7.02	7
Weather condition	4.76	4	4.47	4	4.36	4
Main cause	8.54	11	8.59	11	8.69	11

According to the obtained rankings from Table 3, it can be concluded that the most important variables affecting road freight accidents in all the three years under study are the accident point geometry (straight, horizontal curve, and intersections), road pavement status, and lighting status, respectively.

### 3.4. Exploratory Factor Analysis

It is inevitable to come across with a large number of variables in any study. In order to obtain more accurate analysis data as well as accomplish more scientific and at the same time practical results, researchers have always been attempting to reduce the number of variables and establish a new structure for them. Therefore, Factor Analysis is typically used to achieve this goal. Factor Analysis tries to identify basic variables or factors to explain the correlation pattern among the observed variables. Factor Analysis plays a pivotal role in identifying the hidden variables or factors utilizing the observed variables.

When performing Factor Analysis, first one should be sure whether the available data could be used for the analysis purpose. In other words, it should be determined whether the number of the intended data (sample size and the relationship between variables) is appropriate for Factor Analysis or not. Consequently, in this study the KMO index and Bartlett test were utilized to test the referred hypothesis.

Table 4 displays the results for the KMO index and Bartlett test in the present study. The more the index approximates 1, the more appropriate the intended data will be for Factor Analysis. Similarly, if KMO index is smaller than 0.5, the Factor Analysis results are not appropriate for the intended data and this analysis should not be used to interpret the results. Besides, the sig value obtained from the Bartlett Test is smaller than 5% for all the cases and therefore the assumption that the correlation Matrix is known is rejected.

**Table 4. KMO and Bartlett's test result**

Year	2016	2017	2018
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.562	0.516	0.524
Bartlett's Test of Sphericity	Approx Chi-Square	1201.311	877.590
	df	66	66
	Sig	0.000	0.000

The tables obtained from Factor Analysis consist of two parts. The first part is for special values which determine the factors that are included in the analysis. Those factors whose specific values are smaller than 1 are excluded from the analysis. In this study, the factors 1, 2, 3, 4, and 5 which their special values are larger than 1 are included in the analysis.

The second part indicates the specific values of the factors extracted through the rotation. It should be noticed that in the rotation of all the remaining factors, a proportion of the total changes, which are explained via these 5 factors, is taken as fixed (approximately 60%). Tables 5 to 7 present the specific values of road freight accidents from 2016 to 2018, respectively.

**Table 5. Specific values of vehicle accidents in 2016**

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.892	15.765	15.765	1.892	15.765	15.765
2	1.570	13.084	28.848	1.570	13.084	28.848
3	1.259	10.489	39.338	1.259	10.489	39.338
4	1.168	9.737	49.075	1.168	9.737	49.075
5	1.088	9.064	58.139	1.088	9.064	58.139
6	0.982	8.180	66.319			
7	0.903	7.521	73.840			
8	0.864	7.197	81.038			
9	0.715	5.962	87.000			
10	0.639	5.326	92.325			
11	0.493	4.105	96.430			
12	0.428	3.570	100.000			

Extraction Method: Principal Component Analysis.

**Table 6. Specific values of vehicle accidents in 2017**

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.696	14.134	14.134	1.696	14.134	14.134
2	1.570	13.082	27.216	1.570	13.082	27.216
3	1.266	10.547	37.763	1.266	10.547	37.763
4	1.188	9.896	47.660	1.188	9.896	47.660
5	1.102	9.183	56.842	1.102	9.183	56.842
6	0.994	8.285	65.127			
7	0.924	7.699	72.826			
8	0.891	7.427	80.254			
9	0.743	6.191	86.445			
10	0.646	5.384	91.828			
11	0.557	4.641	96.470			
12	0.424	3.530	100.000			

Extraction Method: Principal Component Analysis.

**Table 7. Specific values of vehicle accidents in 2018**

Component	Total Variance Explained					
	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.746	14.551	14.551	1.746	14.551	14.551
2	1.518	12.654	27.205	1.518	12.654	27.205
3	1.378	11.487	38.692	1.378	11.487	38.692
4	1.246	10.383	49.074	1.246	10.383	49.074
5	1.047	8.728	57.802	1.047	8.728	57.802
6	1.000	8.332	66.134			
7	0.906	7.553	73.687			
8	0.851	7.096	80.782			
9	0.722	6.018	86.801			
10	0.642	5.353	92.154			
11	0.554	4.614	96.768			
12	0.388	3.232	100.000			

Extraction Method: Principal Component Analysis.

Tables 8 to 10 indicate the rotated matrix of the components from 2016 to 2018, which include the factor loads for each variable in the factors remaining after rotation. The higher the absolute value of these coefficients in each row, the more noticeable role the related factor plays in the total changes of the given variable.

**Table 8. Component rotation matrix for vehicle accidents in 2016**

	Component				
	1	2	3	4	5
Time of the accidents	0.136	0.752	0.364	0.066	0.209
Season of the accident	0.673	-0.136	-0.054	0.231	-0.162
Day of the accident	0.053	-0.054	-0.231	0.016	0.661
Type of highway function	0.055	-0.403	0.573	-0.177	0.223
Highway surface condition	0.747	-0.229	-0.104	-0.071	0.081
Accident point geometry	-0.136	-0.076	-0.325	0.418	0.402
Lighting condition	0.390	0.648	0.372	0.100	0.163
Faulty freight vehicle	-0.021	-0.374	0.491	0.470	-0.086
Faulty driver's age	-0.003	-0.137	0.294	0.269	-0.326
Non faulty vehicle	-0.003	0.378	-0.361	0.460	-0.378
Weather condition	0.826	-0.115	-0.158	-0.045	-0.028
Main cause	-0.058	-0.168	0.055	0.618	0.283

**Table 9. Component rotation matrix for vehicle accidents in 2017**

	Component				
	1	2	3	4	5
Time of the accidents	-0.074	0.699	-0.316	0.285	-0.140
Season of the accident	0.297	-0.058	-0.111	0.154	0.460
Day of the accident	-0.146	0.111	-0.069	-0.252	0.344
Type of highway function	0.389	-0.474	-0.171	0.323	-0.397
Highway surface condition	0.824	0.098	0.004	-0.165	0.105
Accident point geometry	0.092	0.124	0.464	0.306	0.331
Lighting condition	0.135	0.654	-0.245	0.404	-0.232
Faulty freight vehicle	0.084	-0.428	0.100	0.642	-0.017
Faulty driver's age	-0.155	0.046	0.036	0.433	0.551
Non faulty vehicle	-0.085	0.402	0.599	-0.033	-0.092
Weather condition	0.804	0.188	-0.044	-0.166	0.136
Main cause	0.198	0.076	0.688	0.052	-0.298

**Table 10. Component rotation matrix for vehicle accidents in 2018**

	Component				
	1	2	3	4	5
Time of the accidents	0.014	-0.625	-0.324	0.405	0.039
Season of the accident	0.288	-0.180	0.266	0.004	-0.036
Day of the accident	-0.120	-0.025	-0.014	0.064	0.825
Type of highway function	0.088	-0.009	0.749	0.245	0.218
Highway surface condition	0.836	-0.015	0.045	-0.214	-0.007
Accident point geometry	0.183	0.465	-0.283	0.182	0.020
Lighting condition	0.274	-0.602	-0.199	0.434	0.142
Faulty freight vehicle	0.026	0.268	0.460	0.598	-0.058
Faulty driver's age	-0.017	-0.022	0.024	0.446	-0.508
Non faulty vehicle	0.175	0.343	-0.539	0.269	0.063
Weather condition	0.844	-0.099	0.056	-0.166	-0.019
Main cause	0.297	0.563	-0.119	0.306	0.174

Extraction Method: Principal Component Analysis.

Based on the Factor Analysis performed for 12 variables affecting the road freight accidents in 2016, as indicated in Table 8, 5 factors are identified as the major factors. Factor Analysis indicates that variables like the season of accident, road pavement condition and weather condition are ranked as the first important category of factors; variables like time of accident and lighting status as the second important category of factors; variables like type of road and faulty vehicle as the third important category of factors; variables like non-faulty vehicle, accident point geometry, and the main cause as the fourth important category of factors; and finally weekdays and the age of the driver are ranked as the fifth important category of factors affecting the severity of road freight accidents in 2016.

Based on the Factor Analysis carried out for 12 variables affecting the road freight accidents in 2017, as indicated in Table 9, 5 factors are identified as major factors. Factor Analysis shows that the first important category of factors include variables related to the road pavement status and weather conditions; while, variables like time of the accident and lighting status are the second most important category of factors; variables like accident point geometry, non-faulty vehicle, and the main cause are the third most important factor; faulty vehicle as the fourth most important factor; and the season of accident, weekdays, type of highway, and age of the driver are ranked as the fifth most important category of factors affecting the severity of road freight accidents in 2017.

Based on the Factor Analysis conducted for 12 variables affecting the road freight accidents in 2018, as presented in Table 10, five factors are identified as major factors. Factor Analysis indicates that factors like season of the accident, road pavement status and weather conditions are ranked as the first most important factor; while variables like time of the accident, lighting status, and main cause as the second most important factor, type of road and non-faulty vehicle as the third; faulty vehicle as the fourth; and weekdays and the age of driver as the fifth most important factor affecting the severity of road freight accidents in 2018.

### 3.5. Road Freight Accidents Severity Model

To establish a model for road freight accident severity, 12 independent variables and 1 dependent variable were defined. Afterward, they were converted into nominal variables (0 and 1) to be aptly used in SPSS. The dependent variable, i.e. accident severity, was defined as injury, fatality, and damage accidents. As there were a small number of fatal accidents, such accidents were categorized as injury accidents. Ultimately, the number of dependent and independent variables was narrowed down into 2 and 12 respectively. The variables include the type of the faulty vehicle, type of the road function, the main cause, type of collision, age of the faulty driver, time of the accident, road pavement status, etc. The enter, backward, and forward methods can be utilized to establish the Logit model. Now, it should be noticed which of the above-mentioned methods contributes to more appropriate output. In other words, one should consider which of the above-mentioned methods can present a better model for road freight accidents in Gilan Province.

To identify such a significant issue, the correct percentage and goodness criteria for the fit model were considered to identify the fitness of the model. The goodness criterion of the fit model is indicated by  $R^2$  parameter. This parameter shows the percentage of the changes in a given dependent variable determined through Logit independent variables. Also, the correct percentage criterion determines to what extent the model prediction is correct. In other words, these two criteria are used to make comparisons between the models and identify a better model. It is performed in such a way that the more the  $R^2$  value approaches unity, a better fit model is established. Similarly, a higher correct percentage value of a given model, indicates a more powerful model in predicting the accidents.

It should be noticed that as in the first method (enter) all the variables simultaneously are input into the equation, the model lacks enough time to appropriately process the data and select the most significant variables; hence, it cannot be an appropriate method. Because of this, the forward and backward methods are used to input the data into Logit equation. Each method which has the highest accuracy in predicting the number of accidents could be adopted as the best method. In the backward and forward methods, respectively, those variables exit and/or enter where by their exit or entrance, the minimum change would occur in the value of  $R^2$  corresponding to the equation. Likewise, exit or entrance of the variable leads to improvement or in other words increase of  $R^2$  value. This method helps us with choosing the way of entering the independent variables to be analyzed. Applying various methods provides us with establishing different equations with the same data and ultimately selecting the best equation.

Table 11 summarizes Logit models in two forward and backward methods. The determiner of the best model is its degree of accuracy when making predictions. Accordingly, the backward method with its higher degree of correct percentage in all the cases is selected as the best method to establish Logit model for the severity of road freight accidents.

**Table 11. Summary of the forward and backward methods**

Logit method	2016	2017	2018
Forward	73.2	74.3	77.7
Backward	82.1	79.9	78.7

As it was explicated, the present backward method model was selected owing to its high degree of accuracy in predicting accidents. Hence, this chapter is devoted to introducing the best model. The Chi-Square statistic is used to determine the effectiveness of dependent variables on independent variables and the fitness of the overall model, which is comparable with F statistics in ordinary regression analysis. Tables 12 to 14 present the backward method model coefficients for the vehicle accidents in 2016, 2017, and 2018, respectively. According to these Tables, the Chi-Square model shows whether the independent variable(s) affects the dependent variable or not. As it is observed, in all the models the Chi-Square values have zero Sig. Therefore, the independent variables affect the dependent variable and indicate a high degree of fitness.

**Table 12. Backward method model coefficients in 2016**

Final Step	Chi-Square	Df	Sig.
Step	8.287	3	0.040
Step 7 Block	630.500	27	0.000
Model	630.500	27	0.000

**Table 13. Backward method model coefficients in 2017**

Final Step	Chi-Square	df	Sig.
Step	17.65	8	0.240
Step 5 Block	518.975	23	0.000
Model	518.975	23	0.000

**Table 14. Backward method model coefficients in 2018**

Final Step	Chi-Square	df	Sig.
Step	7.255	2	0.270
Step 7 Block	630.459	30	0.000
Model	630.459	30	0.000

By selecting the backward method and then entering all the selected variables into the process of model development; and after passing a variety of stages, the ultimate model is obtained through this method. Tables 15 to 17 display the variables entered into the model, using the Wald test, and ultimate Logit model. Wald test considers the significance of the variables entered into the regression equation and it is comparable with t statistics in normal regression. Based on these Tables, the Logit model for different years can be presented in this way; the Positive coefficients of the independent variables indicate their positive relationship with the severity of accidents; while, the negative coefficients demonstrate their negative relation with the severity of accidents.

Based on Tables 15 to 17, in 2016, variables like autumn, daylight time, the drivers aging 18 to 60, pickup truck, and not following safety guidelines in case of freighting goods were the most effective variables on the severity of road freight accidents. Whereas, in 2017, such variables like rural, trucks and trailers, non-faulty private vehicles, motors, pedestrians, and no following safety guidelines were the most effective variables. When considering the most effective variables in 2018, it was determined that the variables of accidents occurring between 12 p.m to 6 p.m, rural, highways, trailers, non-faulty motorcycles, disregarding front are the main causes, and not following safety guidelines when freighting are the effective variables on road freight accidents.

**Table 15. Logit model variables for vehicle accident severity in 2016**

Variables	Beta	S.E.	Wald	Sig.	Exp (Beta)
Season (Autumn)	-0.662	0.255	6.725	0.010	0.516
Daylight (day)	-0.462	0.184	6.300	0.012	0.630
faulty driver's age (18-30)	-1.300	0.409	10.104	0.001	0.273
faulty driver's age (31-45)	-1.371	0.389	12.444	0.000	0.254
faulty driver's age (46-60)	-1.289	0.405	10.117	0.001	0.275
faulty freight vehicle (pickup trucks)	1.087	0.510	0.002	0.046	1.024
not following safety guidelines when freighting	-1.237	0.629	3.862	0.049	0.290
Constant	3.464	0.000	4.591	0.000	3.195

**Table 16. Logit model variables for vehicle accident severity in 2017**

Variables	Beta	S.E.	Wald	Sig.	Exp (Beta)
Rural road	1.500	0.700	4.586	0.032	4.480
Daylight (day)	-0.463	0.185	6.243	0.012	0.630
faulty freight vehicle (trucks)	1.757	0.853	4.246	0.039	5.794
faulty freight vehicle (trailers)	0.149	0.379	0.156	0.043	1.161
non-faulty vehicle (private cars)	-1.625	0.727	5.001	0.025	0.197
non-faulty vehicle (motorcycles)	3.681	1.023	12.935	0.000	9.687
Pedestrians	3.569	1.239	8.294	0.004	5.464
not following safety guidelines when freighting	-1.713	0.633	7.320	0.007	0.180
Constant	1.183	0.879	1.811	0.178	3.265

**Table 17. Logit model variables for vehicle accident severity in 2018**

Variables	Beta	S.E.	Wald	Sig.	Exp (Beta)
Rural road	-0.701	0.206	11.535	0.001	0.496
Daylight (day)	1.385	0.620	4.990	0.025	3.997
faulty freight vehicle (trucks)	0.800	0.399	4.026	0.045	2.225
faulty freight vehicle (trailers)	1.521	0.916	2.754	0.097	4.575
non-faulty vehicle (private cars)	3.877	0.623	3.781	0.000	4.261
non-faulty vehicle (motorcycles)	0.706	0.329	4.602	0.032	0.454
not following safety guidelines when freighting	-0.789	0.516	2.341	0.026	0.454
Constant	-0.379	2.627	0.000	1.000	0.684

The most noticeable points obtained from the type of accidents over the years in this case study reveal the need for special and more attention of the authorities responsible for road accidents. These include the police and ministry of road and city planning, rural planners, trailers, motorists, as well as drivers' following safety guidelines while freighting loads, all of these variables have been effective on the road freight accidents over the years 2016, 2017, and 2018.

#### 4. Conclusion

Through probing into the data concerning the road freight accidents, the present article focused on analyzing the effects of various variables on the severity of accidents from 2016 to 2018. Based on the results of the frequency of



variables, it was determined that more than 50% and 40% of road freight accidents were related to the damage and injury accidents, respectively. The accidents leading to fatality just constituted 10% of the total number of accidents. The majority of accidents occur between 6 a.m to 12 p.m and 12 p.m to 6 p.m on weekdays in autumn. In more than 70% of the cases, road pavement is dry, the weather is sunny, the route is straight, and it is daytime. Among the faulty freight vehicles in road accidents, the pickup trucks and trucks are engaged in 60% and 30% of the accidents, respectively. Other freight vehicles account for 10% of road the freight accidents. Furthermore, according to the obtained ranking, variables of accident point geometry (tangent, horizontal curve, and intersections), road pavement status, and lighting status have the highest influence on the road freight accident severity.

By performing the Factor Analysis method, 5 most effective factors on road accidents in 2016, 2017, and 2018 were identified. Moreover, the output of the results obtained from Logit model indicates that the severity of accidents increases as the result of trailers and motorcycles presence. Moreover, regarding freighting goods in rural areas, some special measures should be taken. Unfortunately, the main cause of road freight accidents is not following safety guidelines by the drivers of freighting vehicles. Hence, some strict measures should be adopted as well as more effective fining strategies should be applied for such drivers.

## 5. Conflicts of Interest

The authors declare no conflict of interest.

## 6. References

- [1] Loprencipe, Giuseppe, Laura Moretti, Giuseppe Cantisani, and Paolo Minati. "Prioritization Methodology for Roadside and Guardrail Improvement: Quantitative Calculation of Safety Level and Optimization of Resources Allocation." *Journal of Traffic and Transportation Engineering (English Edition)* 5, no. 5 (October 2018): 348–360. doi:10.1016/j.jtte.2018.03.004.
- [2] Xu, Pengpeng, Helai Huang, and Ni Dong. "The Modifiable Areal Unit Problem in Traffic Safety: Basic Issue, Potential Solutions and Future Research." *Journal of Traffic and Transportation Engineering (English Edition)* 5, no. 1 (February 2018): 73–82. doi:10.1016/j.jtte.2015.09.010.
- [3] Lin, Pei-Sung, Rui Guo, Elzbieta Bialkowska-Jelinska, Achilleas Kourtellis, and Yu Zhang. "Development of Countermeasures to Effectively Improve Pedestrian Safety in Low-Income Areas." *Journal of Traffic and Transportation Engineering (English Edition)* 6, no. 2 (April 2019): 162–174. doi:10.1016/j.jtte.2019.02.001.
- [4] Ihueze, Chukwutoo C., and Uchendu O. Onwurah. "Road Traffic Accidents Prediction Modelling: An Analysis of Anambra State, Nigeria." *Accident Analysis & Prevention* 112 (March 2018): 21–29. doi:10.1016/j.aap.2017.12.016.
- [5] Kiran, Bangaram Naga, Nekkanti Kumaraswamy, and Chundupalli Sashidhar. "A review of road crash prediction models for developed countries." *American journal of traffic and transportation engineering* 2, no. 2 (2017): 10-25.
- [6] Abdulhafedh, Azad. "Road Crash Prediction Models: Different Statistical Modeling Approaches." *Journal of Transportation Technologies* 07, no. 02 (2017): 190–205. doi:10.4236/jtts.2017.72014.
- [7] Cai, Qing, Mohamed Abdel-Aty, Jaeyoung Lee, and Naveen Eluru. "Comparative Analysis of Zonal Systems for Macro-Level Crash Modeling." *Journal of Safety Research* 61 (June 2017): 157–166. doi:10.1016/j.jsr.2017.02.018.
- [8] Zhong-xiang, Feng, Lu Shi-sheng, Zhang Wei-hua, and Zhang Nan-nan. "Combined Prediction Model of Death Toll for Road Traffic Accidents Based on Independent and Dependent Variables." *Computational Intelligence and Neuroscience* 2014 (2014): 1–7. doi:10.1155/2014/103196.
- [9] Renne, John L., and Tara Tolford. "A Planning Tool for Evaluating Vehicles Miles Travelled and Traffic Safety Forecasts of Growth Management Scenarios: A Case Study of Baton Rouge and New Orleans." *Transportation Research Part D: Transport and Environment* 59 (March 2018): 237–245. doi:10.1016/j.trd.2018.01.001.
- [10] Yasmin, Shamsunnahar, Salah Uddin Momtaz, Tammam Nashad, and Naveen Eluru. "A Multivariate Copula-Based Macro-Level Crash Count Model." *Transportation Research Record: Journal of the Transportation Research Board* 2672, no. 30 (October 6, 2018): 64–75. doi:10.1177/0361198118801348.
- [11] Ziari, Hasan, Amir Amini, Amir Saadatjoo, Sayyed Mohsen Hosseini, and Vahid Najafi Moghaddam Gilani. "A Prioritization Model for the Immunization of Accident Prone Using Multi-criteria Decision Methods and Fuzzy Hierarchy Algorithm." *Computational Research Progress in Applied Science & Engineering (CRPASE)* 3, no. 3 (2017).
- [12] Bargegol, Iraj, Vahid Najafi Moghaddam Gilani. "Modelling of urban accidents using logistic regression." *The 1th National Conference on Highway and Transportation, University of Guilan, Rasht, Iran* (2017).
- [13] Memon, Abdul Qadeer. "Modelling road accidents from national datasets: a case study of Great Britain." PhD diss., UCL (University College London), (2012).

- [14] Bahadorimonfared, Ayad, Hamid Soori, Yadollah Mehrabi, Ali Delpisheh, Alireza Esmaili, Masoud Salehi, and Mahmood Bakhtiyari. "Trends of Fatal Road Traffic Injuries in Iran (2004–2011)." Edited by Martyn Kirk. PLoS ONE 8, no. 5 (May 28, 2013): e65198. doi:10.1371/journal.pone.0065198.
- [15] Ghaffar, Abdul, Adnan A Hyder, and Tayyeb I Masud. "The Burden of Road Traffic Injuries in Developing Countries: The 1st National Injury Survey of Pakistan." Public Health 118, no. 3 (April 2004): 211–217. doi:10.1016/j.puhe.2003.05.003.
- [16] Labinjo, M, C Juillard, O C Kobusingye, and A A Hyder. "The Burden of Road Traffic Injuries in Nigeria: Results of a Population-Based Survey." Injury Prevention 15, no. 3 (June 1, 2009): 157–162. doi:10.1136/ip.2008.020255.
- [17] Hu, Sitao, and Qiaojun Xiang. "Characteristics Analysis of Traffic Accidents on Rural Roads." CICTP 2012 (July 23, 2012): 2506-2513. doi:10.1061/9780784412442.255.
- [18] Zangooei Dovom, Hossein, Yousef Shafahi, and Mehdi Zangooei Dovom. "Fatal Accident Distribution by Age, Gender and Head Injury, and Death Probability at Accident Scene in Mashhad, Iran, 2006–2009." International Journal of Injury Control and Safety Promotion 20, no. 2 (June 2013): 121–133. doi:10.1080/17457300.2012.692694.
- [19] Lee, Sangbok, and Byung Yong Jeong. "Comparisons of Traffic Collisions between Expressways and Rural Roads in Truck Drivers." Safety and Health at Work 7, no. 1 (March 2016): 38–42. doi:10.1016/j.shaw.2015.11.005.
- [20] Behbahani, Hamid, Mehdi Jahangir Samet, Vahid Najafi Moghaddam Gilani, and Amir Amini. "Determining of the Parking Manoeuvre and the Taxi Blockage Adjustment Factor for the Saturation Flow Rate at the Outlet Legs of Signalized Intersections: Case Study from Rasht City (Iran)." IOP Conference Series: Materials Science and Engineering 245 (October 2017): 042017. doi:10.1088/1757-899x/245/4/042017.
- [21] Bargegol, Iraj, Mohammad Nikookar, Reza Vatani Nezafat, Esmat Jafarpour Lashkani, and Arash M. Roshandeh. "Timing Optimization of Signalized Intersections Using Shockwave Theory by Genetic Algorithm." Computational Research Progress in Applied Science & Engineering (CRPASE) 1, no. 4 (2015).
- [22] Iranian Road Transport Statistics. Ministry of Roads and Urban Development of the Islamic Republic of Iran, (2019).
- [23] Bargegol, Iraj, Vahid Najafi Moghaddam Gilani, and Sadra Farghedayn. "Analysis of the effect of vehicles conflict on pedestrian's crossing speed in signalized and un-signalized intersection." Advances in Environmental Biology (2014): 502-510.
- [24] Chatfield, Mark, and Adrian Mander. "The Skillings–Mack Test (Friedman Test When There Are Missing Data)." The Stata Journal: Promoting Communications on Statistics and Stata 9, no. 2 (August 2009): 299–305. doi:10.1177/1536867x0900900208.
- [25] Thompson, B., 2004. Exploratory and confirmatory factor analysis: Understanding concepts and applications. American Psychological Association. Young, Rhonda Kae, and Joel Liesman. "Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model." Accident Analysis & Prevention 39, no. 3 (2007): 574-580.
- [26] Abdi, Ali, Omid Nassimi, Reza Salehfard, and Vahid Najafi Moghaddam. "Analysing the Influence of Encroachment Angle and Median Parameters on Safety of Rural Highways Using Vehicle Dynamics Performance." IOP Conference Series: Materials Science and Engineering 471 (February 23, 2019): 062043. doi:10.1088/1757-899x/471/6/062043.
- [27] Jamshidpour, Faramarz, and Barg E. Gol. "Relationship between Pedestrians' Speed, Density and Flow Rate of Crossings through Urban Intersections (Case Study: Rasht Metropolis) (Research Note)." International Journal of Engineering 30, no. 12 (2017): 1814-1821.