International Journal of Applied Operational Research
Vol. 5, No. 3, pp. 59-69, Summer 2015

Journal homepage: ijorlu.liau.ac.ir

# Multivariate Statistical Analysis Decision-making Hybrid Method for Road Traffic Safety Evaluation in Iran

A. M. Rahimi, F. Movahedi Sobhani, S. Mirzaei Khezri*

**Received:** 14 November 2014 ; **Accepted:** 19 April 2015

**Abstract** Obviously, improving the road safety and the efficient allocation of limited resources to the provinces according to their ranking should be done. This paper presents a hybrid method of multivariate statistical analysis-decision making to evaluate Iran road traffic safety. In order to solve the problems of road traffic safety, a macroscopic evaluation and traffic safety level classification in Iran was carried out. An index system which consists of 14 relative indexes for road traffic safety evaluation was established. The principal component analysis method was used to reduce the dimensions of the multi-index data. Based on this, 2 components were extracted. The Index of Road Traffic Safety (IRTS) was calculated to rank the provinces of the country. A K-means method was applied to classify the provinces. A TOPSIS technique was used to examine the status of each cluster in terms of safety levels. Results showed that there are 4 safety levels entitled good, average, weak and very weak. The levels are approximately similar to result of the rankings.

**Keywords:** principal component analysis, K-means, TOPSIS , Index of Road Traffic Safety.

## 1 Introduction

Road traffic accidents are one of the most prominent public health threats in the world [1]. According to statistics, each year, 1.2 million people die and 50 million people are injured in road accidents around the world [2]. Iran has one of the highest rates of traffic accident fatalities and injuries. Thus, road safety is one of the main concerns of the Iranian transportation industry and a great deal of expenditure is incurred to control road traffic accidents. This article attempts to recognize and rank the country's provinces that provides good information for transportation planners in terms of the following criteria: death rate of per 100,000 people, injury rate of per 100,000 people, death rate of per 100 kilometers, injury rate of per 100 kilometers, lighting rate per 100 kilometers, speed camera rate of per 100

* **Corresponding Author.** (✉)
**E-mail:** so.mirzaeek@hotmail.com (S.Mirzaei Khezri)

**A. M. Rahimi**
Assistant Professor, Department of Civil Engineering, University of Zanjan, Zanjan, Iran.

**F. Movahedi Sobhani**
Assistant Professor, Department of Industrial Engineering, Science and Research branch, Islamic Azad University, Tehran, Iran.

**S. Mirzaei Khezri**
Ph.D. Candidate, Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

kilometers, death rate to daily traffic volume, injury rate to daily traffic volume, emergency bases rate of per 100,000 people, crescent bases rate of per 100,000 people, emergency bases rate of per 100 kilometers, crescent bases rate of per 100 kilometers, average hours of driver training and the number of active police stations. Obviously, improving road safety and the efficient allocation of limited resources to the provinces according to their ranking is done. For this purpose, an indexing system which consists of 14 relative indexes for the road traffic safety evaluation was established. The principal component analysis method was used to reduce the dimensions of the multi-index data. The Index of Road Traffic Safety (IRTS) was produced in order to rank the road traffic safety situation. The K-means clustering method was applied to classify provinces taking two principal components. Finally, multi-criteria decision-making techniques such as TOPSIS were implemented in order to examine the status of each cluster in terms of safety levels. One of the main advantages of this paper is the combination of multi-criteria decision-making and multivariate statistical analysis techniques in the ranking of provinces.

The current paper is outlined as this. Section 2 contains a review of the previous studies conducted on ranking. In Section 3, the methodology and introduction to the principal component analysis (PCA)-cluster analysis(CA) method and the TOPSIS technique are discussed. In Section 4, the data used is described. Analysis of the results is given in section 5. Finally, the conclusion to our study is presented in Section 6.

## 2 Literature review

Fancello et al. [3],who examined road networks of suburban Sardinia, Italy, using a decision support system based on Electre III. This means that the main route is divided into 10 homogeneous segments to assess the safety with indicators such as peak hours, the percentage of heavy vehicles and accident rate. Khorasani et al. [4], evaluated the safety performance of 21 European countries. These are analyzed using decision-making methods such as SAW, AHP and Fuzzy TOPSIS. For this purpose, 11 indicators including the following were defined: the rate of the use of seat belts for front and rear seats, the percentage of cars with an average age of over 6 years, the cost of health care as a share of GDP, the average age of the passenger fleet, the share of motorcycles in the fleet, the volume of heavy vehicles in the fleet and the share of motorcycles in total road network.

Lin et al. [5],who examined China's road traffic safety in macroscopic indexes which consisted of 14 relative indexes. These indexes were established having taken into account the human, vehicle, road, and socio-economic synthetic influences by principal component analysis-cluster analysis method. Ahmadvand et al. [6], evaluated the performance of their selected provinces' road safety by applying model of data envelopment analysis (DEA) with principal component analysis. These focused on input variables such as age of the fleet, the percentage of roads with lighting, the percentage of highways, the elimination of black spots, the number of police stations in each province and the percentage of train users.

Furthermore, the output variables in the number of accidents and deaths were also investigated. The road safety situation of the Bushehr province is investigated by Haghighat [7]. In this study, all measures affecting road safety standards of the Bushehr province were categorized using the group analytic hierarchy process (GAHP). Following this, the roads of the Bushehr province were ranked using TOPSIS. Wei and Sun [8],who investigated the road safety of the eastern provinces of China with on the basis of improved principal component analysis and cluster analysis. Molla et al. [9], identified the principal components and factors

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

associated with road traffic crashes in the U.S. through a retrospective review based on more than two million records of fatal crashes over the space of 38 years (1975-2012). This information was taken from the National Highway Traffic Safety Administration Official's Fatal Accident Reporting System (FARS) database.

## 3 Methodology

In this section, we briefly explain the principal component analysis method that was used to reduce the dimensions of the multi-index. The K-means Clustering Analysis and TOPSIS methods that followed were described.

### 3.1 Principal component analysis

Principal component analysis (PCA) method was used to reduce dimensions of the multi-index data and extract principal components, and then, the Index of Road Traffic Safety (IRTS) was constructed for road traffic safety situation.

Suppose that $X_{ij}$ is the value of the No. j index of No. i sample ($i=1,2,3,...,n$ ; $j=1,2,3,...,p$ ; n is the number of samples; P is the number of indexes) , the steps of the PCA method are shown as follows[10]:

**Step 1**: Standardization of original index data In order to eliminate the influence of the order of magnitude and the dimensional effect on evaluation results, standardize the original index data according to Equation

$$X_{ij} = (x_{ij} - \overline{x}_j) / s_j \quad (1)$$

Where:

$\overline{x}_j$ = Mean of the No. j sample index.

$s_j$ = Standard deviation of the No. j sample index.

**Step 2**: Calculation index correlation matrix.

$$R = (r_{jk})_{p\times p} \qquad j=1,2,3,...,p \quad ; \quad k=1,2,3,...,p \quad (2)$$

Where:

$r_{jk}$ =The correlation coefficient of the No. j and the No. k indexes, can be computed as follows

$$r_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(Z_{ij}Z_{ik}) \quad \text{and} \quad r_{ii}=1, r_{jk}=r_{kj} \quad (3)$$

**Step 3**: Calculation eigenvalues and eigenvectors

Solve the eigen equation $|\lambda_p - R| = 0$ to obtain the eigenvalues $\lambda_g$ ($g=1,2,...,p$), then, arrange $\lambda_g$

in order of size: $\lambda_1 \geq \lambda_2 .... \geq \lambda_p \geq 0$, the corresponding eigenvectors are $L_g = (L_{g_1}, L_{g_2}, ..., L_{g_p})$

**Step 4**: Calculate distribution rate and ascertain the number of principal components

Distribution rate of principal components is defined as:

$$\lambda_g / \sum_{g=1}^{p}\lambda_g \quad (4)$$

Cumulative distribution rate of the previous principal components is:

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

$$\sum_{i=1}^{k}\lambda_i / \sum_{g=1}^{p}\lambda_g \tag{5}$$

**Step 5**: Construction Index of Road Traffic Safety (IRTS)

$$IRTS = \sum_{g=1}^{k}\left[\left[\lambda_g / \sum_{g=1}^{p}\lambda_g\right]F_g\right] \tag{6}$$

Where:

$F_g$ = The score of the No. g principal component can be calculated as follows[10]:

$$F_g = l_{g1}Z_1 + l_{g2}Z_2 + \ldots + l_{gp}Z_p \qquad (g = 1, 2, \ldots, k) \tag{7}$$

The K-means clustering method was applied to safety level classification taking IRTS as variables.

### 3.2 K-means

The K-means algorithm is the simplest and most commonly algorithmis based on the square-error criterion. The general objective is to obtain the partition that, for a fixed number of clusters, minimizes the total square-error.

It starts with a random, initial partition and keeps reassigning the samples to clusters, based on the similarity between samples and clusters, until a convergence criterion is met. Typically, this criterion is met when there is no reassignment of any sample from one cluster to another that will cause a decrease of the total squared error [10].

Suppose that the given set of N samples in an n-dimensional space has somehow been partitioned into K clusters $\{C_1, C_2, \ldots C_K\}$. Each $C_K$ has $n_k$ samples and each sample is in exactly one cluster, so that $\sum n_k = N$, where k = 1, . . . , K. The mean vector $M_K$ of cluster $C_K$ is defined as the centroid of the cluster or

$$M_K = (1/n_k)\sum_{i=1}^{n_k} x_{ik} \tag{8}$$

Where $x_{ik}$ is the ith sample belonging to cluster $C_K$. The square-error for cluster $C_K$ is the sum of the squared Euclidean distances between each sample in $C_K$ and its centroid. This error is also called the within - cluster variation:

$$e_k^2 = \sum_{i=1}^{n_k}(x_{ik} - M_K)^2 \tag{9}$$

The square - error for the entire clustering space containing K clusters is the sum of the within - cluster variations:

$$E_k^2 = \sum_{k=1}^{k} e_k^2 \tag{10}$$

The steps of K-means algorithm are shown as follows:

**Step1**: Select an initial partition with K clusters containing randomly chosen samples, and compute the centroids of the clusters.

**Step2**: Generate a new partition by assigning each sample to the closest cluster center.

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

**Step3**: Compute new cluster centers as the centroids of the clusters.
**Step4**: Repeat steps 2 and 3 until an optimum value of the criterion function is found (or until the cluster membership stabilizes) [10].

### 3.3 TOPSIS

The TOPSIS was first developed by Hwang & Yoon. According to this technique, the best alternative would be the one that is nearest to the positive-ideal solution and farthest from the negative ideal solution [11]. The calculation processes of the method are as following:
The first step is to convert the decision matrix $R=\left[r_{ij}\right]$ in a matrix of scale by using the following equation. If m is number of options and n is the number of criteria, and $N=\left[n_{ij}\right]_{m\times n}$ So, we have following equation:

$$n_{ij}=\frac{r_{ij}}{\sqrt{\sum_{i=1}^{m}(r_{ij})^2}} \tag{11}$$

**Step 2**: Calculate the weighted normalized decision matrix. The weighted normalized value $v_{ij}$ is calculated as :

$$v_{ij}=w_j\times n_{ij} \quad ,i=1,...,m, \quad j=1,...,n \tag{12}$$

Where $w_j$ is the weight if the ith criterion, and $\sum_{j=1}^{n} w_j=1$.

**Step 3**: Determine the positive ideal solutions and negative ideal solutions respectively

$$A^+=\left\{\left[\left[\max_i V_{ij}\middle| j\in J_1\right],\left[\min_i V_{ij}\middle| j\in J_2\right]\right]\middle| i=1,2,...,m\right\}=\left\{V_1^+,...,V_m^+\right\} \tag{13}$$

$$A^-=\left\{\left[\left[\min_i V_{ij}\middle| j\in J_1\right],\left[\max_i V_{ij}\middle| j\in J_2\right]\right]\middle| i=1,2,...,m\right\}=\left\{V_1^-,...,V_m^-\right\} \tag{14}$$

Where $J_1$ is associated with the positive criteria, and $J_2$ is associated with the negative criteria
**Step 4**: Calculate the separation measures using the n-dimensional Euclidean distance. The separation of each alternative from the ideal solution is given as:

$$d_i^+=\sqrt{\sum_{j=1}^{n}(v_{ij}-v_j^+)^2} \tag{15}$$

Similarly, the separation from the negative-ideal solution is given as

$$d_i^-=\sqrt{\sum_{j=1}^{n}(v_{ij}-v_j^-)^2} \tag{16}$$

**Step 5**: Calculate the relative closeness to the ideal solution. The relative closeness of the alternative $A_i$ with respect to $A^+$ is defined as

$$CL=\frac{d_i^-}{d_i^-+d_i^+} \tag{17}$$

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

**Step6**: Rank the preference order. A large value of closeness coefficient indicates a good performance of the alternative $A_i$ .The best alternative is the one with the greatest relative closeness to the ideal Solution [12].

## 4 Data

In this section, before the ranking the provinces, road safety measures have been identified. Based on data from the Road Maintenance & Transportation Organization (RMTO), 14 indicators of road safety were defined. This was so the indicators show the road safety of the provinces. Indicators are shown in Table 1.

**Table 1** Indicators of road safety

| Index | Definition |
|---|---|
| S1 | Death rate of per 100,000 persons |
| S2 | Injury rate of per 100,000 persons |
| S3 | Death rate of per 100 kilometers |
| S4 | Injury rate of per 100 kilometers |
| S5 | Lighting rate of per 100 kilometers |
| S6 | Speed Camera rate of per 100 kilometers |
| S7 | Death rate to daily traffic volume |
| S8 | Injury rate to daily traffic volume |
| S9 | Emergency Base rate of 100,000 persons |
| S10 | Crescent Base rate of 100,000 persons |
| S11 | Emergency Base rate of 100 kilometers |
| S12 | Crescent Base rate of 100 kilometers |
| S13 | Average hours of training drivers |
| S14 | The number of active police stations |

## 5 Analysis of results

In this section, the results of the principal component analysis, cluster analysis and TOPSIS method were explored. In the presented research, the traffic accidents data of Iran's provinces is selected from 2013[13]. R software was used to conduct the principal component analysis and cluster analysis, then in order to execute the TOPSIS model; MATLAB 2012 Ra software was used.

### 5.1 Number of principal components to extract

Table 2 indicates that the first, second, third and forth principal components account for 42.6%, 17.5%, 10.4% and 8.8% of total variance respectively.

**Table 2** Results of principal component analysis

| | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 | F11 | F12 | F13 | F14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard Deviation | 2.441 | 1.564 | 1.205 | 1.109 | 0.911 | 0.859 | 0.714 | 0.578 | 0.449 | 0.423 | 0.267 | 0.189 | 0.115 | 0.020 |
| Proportion of Variance | 0.426 | 0.175 | 0.104 | 0.088 | 0.059 | 0.053 | 0.036 | 0.024 | 0.014 | 0.013 | 0.005 | 0.003 | 0.001 | 0.000 |
| Cumulative | 0.426 | 0.600 | 0.704 | 0.792 | 0.851 | 0.904 | 0.940 | 0.964 | 0.979 | 0.991 | 0.996 | 0.999 | 1.000 | 1.000 |

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

In this paper, in order to determine the number of the principal components of the combination, a scree plot and parallel was used. The scree plot, proposed by Cattell, is very popular. In this rule, a plot of the eigenvalues against the number of components forms an "elbow". The number of principal components that need to be retained is shown by the elbow. In many instances, the scree plot may be so smooth that it may not be possible to determine a clear elbow. Hom has suggested a procedure, called parallel analysis. Figure 1 illustrates the two principal components that should be extracted. The first and second principal components account for 42.6% and 17.5% of the variance respectively. The cumulative distribution rate of the previous two principal components reached 60%.

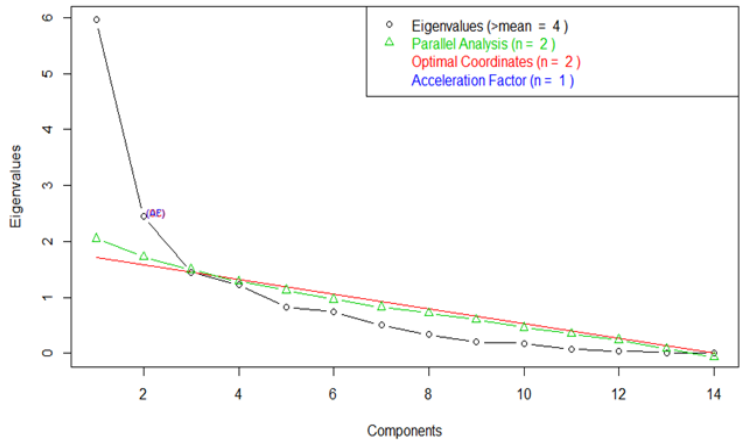



**Fig.1** Scree plot and plot of eigenvalues from parallel analysis

## 5.2 Interpreting principal components

Since the principal components are linear combinations of the original variables, it is often necessary to interpret or provide a meaning for the linear combination. The higher the loading of a variable, the more influence it has in the formation of the principal component score and vice versa. Therefore, we can use the loadings to determine which variables are influential in the formation of the principal components, and we can then assign a meaning or label to the principal component. In this paper, the loading value of .65 has been used as the cutoff point. As shown in Table 3, loading values which are more than .65 were highlighted. It can be said that the first principal component (PC1) represents the road index, and the second principal component represents the monitoring and control index. In other words, the first principal component is a measure of the index of death rate of per 100 kilometers, injury rate of per 100 kilometers, lighting rate of per 100 kilometers, speed cameras rate per 100 kilometers, emergency bases rate of per 100 kilometers and crescent bases rate of per 100 kilometers across the provinces. The second principal component (PC2) is a measure of the index of injury rate to daily traffic volume and the number of active police bases. Therefore, PC1 can be labeled as the road index and PC2 as the monitoring and control index. The principal components' scores are shown in Table 4.

**Table 3** Loading values

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 | S12 | S13 | S14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PCA1 | 0.52 | 0.09 | -0.92 | -0.90 | -0.92 | -0.76 | 0.55 | 0.54 | 0.61 | 0.48 | -0.88 | -0.76 | 0.17 | 0.145 |
| PCA2 | 0.42 | 0.38 | -0.17 | -0.19 | 0.10 | 0.20 | -0.53 | -0.67 | 0.52 | 0.63 | -0.02 | 0.22 | -0.22 | -0.73 |

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

**Table 4** Principal components scores

| Province | PCA1 | PCA2 | Province | PCA1 | PCA2 |
|---|---|---|---|---|---|
| Sharghi | 0.119699 | -1.51199 | Qazvin | -0.781 | 1.513431 |
| Gharbi | 0.45007 | -1.34437 | Qom | -3.44641 | 1.79814 |
| Ardebil | -0.24543 | 0.348708 | Kordestan | 0.741826 | -0.46933 |
| Esfehan | 0.620427 | -1.72697 | Kerman | 1.794277 | -1.40594 |
| Alborz | -7.50009 | 0.590583 | Kermanshah | 0.955223 | -0.93271 |
| Eilam | 1.547405 | 2.464702 | Kohkiloye | 1.448702 | 0.445264 |
| Boshehr | 0.809103 | 0.894224 | Golestan | -1.12893 | 0.435193 |
| Tehran | -7.76818 | -2.01037 | Gilan | -0.77332 | -0.30495 |
| Charmahal | 0.722998 | 0.913938 | Lorestan | -0.03438 | -0.39358 |
| Khorasan joonobi | 3.811955 | 2.540165 | Mazandaran | -1.42206 | 0.085742 |
| Khorasan razvi | 1.149812 | -2.87296 | Markazi | 0.343605 | 0.72947 |
| Khorasan shomali | -0.1953 | 1.65435 | | | |
| Khozestan | 1.061681 | -2.20977 | Hormozgan | 0.293537 | 0.462625 |
| Zanjan | 0.299977 | 1.051823 | | | |
| Cemnan | 1.721491 | 2.676066 | Hamedan | -0.15355 | 0.396849 |
| Sistan va Balochestan | 3.497397 | -3.26291 | | | |
| Fars | 1.276283 | -1.40849 | Yazd | 0.783183 | 0.853083 |

## 5.3 Index of Road Traffic Safety (IRTS)

The scores of the principal components (road and monitoring and control) are used to create the index of road traffic safety that is shown in Table 5. Based on the safety index, the provinces can be ranked. The greater the IRTS, the worse the road traffic safety condition will be.

**Table 5** Result of provinces ranking

| Province | Rank | IRTS | Province | Rank | IRTS | Province | Rank | IRTS |
|---|---|---|---|---|---|---|---|---|
| Tehran | 1 | -3.6573 | Esfehan | 12 | -0.0376 | Charmahal | 23 | 0.4674 |
| Alborz | 2 | -3.0889 | Khorasan razvi | 13 | -0.0125 | Yazd | 24 | 0.4824 |
| Qom | 3 | -1.1527 | Hamedan | 14 | 0.0040 | Boshehr | 25 | 0.5006 |
| Mazandaran | 4 | -0.5902 | Khozestan | 15 | 0.0658 | Kerman | 26 | 0.5180 |
| Golestan | 5 | -0.4044 | Hormozgan | 16 | 0.2057 | Kohkiloye | 27 | 0.6944 |
| Gilan | 6 | -0.3824 | Khorasan shomali | 17 | 0.2059 | Sistan va Balochestan | 28 | 0.9185 |
| Azarbayjan Sharghi | 7 | -0.2132 | Kordestan | 18 | 0.2337 | Eilam | 29 | 1.0892 |
| Lorestan | 8 | -0.0834 | Kermanshah | 19 | 0.2436 | Cemnan | 30 | 1.2002 |
| Qazvin | 9 | -0.0680 | Markazi | 20 | 0.2737 | | | |
| Ardebil | 10 | -0.0435 | Fars | 21 | 0.2971 | Khorasan Joonobi | 31 | 2.0661 |
| Azarbayjan Gharbi | 11 | -0.0433 | Zanjan | 22 | 0.3114 | | | |

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

## 5.4 Clustering provinces in terms of safety

In this section, the principal components' scores (road and monitoring and control) were used to cluster the provinces with the use of the K-means method. For this purpose, the number of clusters must first be determined. The number of clusters is determined based on within groups sum of square. For this purpose, Figure 2 plots of the sum of the squares within the cluster against the number of components and is examined for an "elbow". According to Figure 2, when the number of clusters were increasing, the sum of the squares within the cluster were decreasing. Figure 2 shows the optimal number of clusters as 4. Four clusters are shown in table 6.

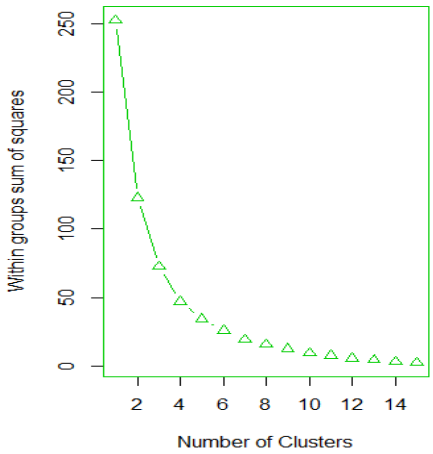



**Fig. 2** Number of cluster vs within groups sum of squares

**Table 6** Provinces clustering

| Cluster | Province |
|---|---|
| 1 | Tehran- Alborz |
| 2 | Eilam- Cemnan- Khorasan Joonobi |
| 3 | Azarbayjan Sharghi- Azarbayjan Gharbi- Esfehan- Khorasan Razvi- Khozestan- Kordestan- Kermanshah- Fars- Kerman- Sistan va Balochestan |
| 4 | Qom- Mazandaran- Golestan- Gilan- Lorestan- Qazvin- Ardebil- hamedan- Hormozgan- Khorasan Shomali- Markazi- Zanjan- Charmahal- Yazd- Boshehr- Kohkiloye |

K-means clustering was applied to the safety level classifications of the provinces. In accordance with the K-means clustering method, road traffic safety was divided into four levels. In this paper, to examine the status of each cluster in terms of safety levels, average values for each index in different clusters were calculated. These can be seen in table 7. The weight of each index was then extracted based on a Shannon Entropy method. Statistics in relation to weight are provided in table 8. The status of safety levels was determined using a TOPSIS model and the results of model are shown in Figure 3.

**Table 7** Average values for each index in different clusters

| | s1 | s2 | s3 | s4 | s5 | s6 | s7 | s8 | s9 | s10 | s11 | s12 | s13 | s14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cluster1 | 13.52 | 314.39 | 110.13 | 2627.13 | 24.52 | 3.33 | 0.01 | 0.21 | 0.68 | 0.16 | 5.09 | 1.20 | 3.57 | 7.00 |
| Cluster2 | 39.10 | 518.86 | 11.38 | 156.91 | 3.04 | 0.00 | 0.05 | 0.62 | 5.30 | 2.12 | 1.50 | 0.59 | 3.69 | 6.00 |
| Cluster3 | 25.80 | 413.57 | 19.77 | 339.85 | 3.06 | 0.20 | 0.07 | 0.97 | 1.87 | 0.52 | 1.40 | 0.40 | 4.66 | 10.60 |
| Cluster4 | 27.66 | 470.36 | 22.52 | 396.39 | 7.20 | 0.90 | 0.02 | 0.43 | 2.20 | 0.85 | 1.59 | 0.60 | 4.58 | 5.76 |

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

**Table 8** Weights of each index in Shannon Entropy

| | | | |
|---|---|---|---|
| w1 | 0.10% | w8 | 12.84% |
| w2 | 0.34% | w9 | 0.53% |
| w3 | 0.35% | w10 | 4.71% |
| w4 | 0.49% | w11 | 0.34% |
| w5 | 13.95% | w12 | 3.17% |
| w6 | 36.85% | w13 | 0.01% |
| w7 | 18.97% | w14 | 7.35% |




**Fig. 3** Result of clusters ranking in TOPSIS method

As can be seen in Figure 3, the first cluster, which includes the Tehran and Alborz provinces, has the highest TOPSIS value of 0.9. These can therefore be labeled as having a good level in terms of safety. The fourth cluster, which includes the Qom, Mazandaran, Golestan, Gilan, Lorestan, Qazvin, Ardebil, Hamedan, Hormozgan, Khorasan Shomali, Markazi, Zanjan, Charmahal, Yazd, Boshehr and Kohkiloye provinces has the second highest TOPSIS value of 0.35. The fourth cluster is much different from the first cluster and is labeled as having an average level of safety. The second cluster, which includes the Eilam, Cemnan and Khorasan Joonobi provinces, has TOPSIS value 0.16 – the third highest. This is therefore labeled as having a weak level of safety. As a result, there are four levels that can be used as labels: good, average, poor and very poor. Results of the methods are provided in Table 9. We can infer that the results of the hybrid combination of K-means classification and TOPSIS with ranking in terms of IRTS are similar to one another.

## 6 Conclusion

In the present study, an index system was defined which consists of 14 relative indexes for road traffic safety evaluation. Using a PCA method to reduced the dimensions of the multi-index data to 2 components. The Index of Road Traffic Safety (IRTS) was constructed for the ranking of the road traffic safety situation. In the second section, the K-means clustering method was applied in order to classify provinces by taking 2 principal components and splitting them into 4 clusters. Then, through the use of a TOPSIS technique, the status of each cluster was examined in terms of safety levels. Results of the hybrid combination of model k-means classification and TOPSIS with ranking by IRTS are similar to one another. In this way, the first cluster, which included the Tehran and Alborz provinces, has a good status in

[ Downloaded from mail.ijaor.ir on 2026-11-27 ]

safety. The fourth cluster, which included the Qom, Mazandaran, Golestan, Gilan, Lorestan, Qazvin, Ardebil, Hamedan, Hormozgan, Khorasan Shomali, Markazi, Zanjan, Charmahal, Yazd, Boshehr and Kohkiloye provinces has an average status in safety due to the fact that the TOPSIS value is less than the TOPSIS value of the first cluster. The second cluster, which included the Eilam, Cemnan and Khorasan Joonobi provinces, was weak in safety. Finally, the third cluster, which included the Azarbayjan Sharghi, Azarbayjan Gharbi, Esfehan, Khorasan Razvi, Khozestan, Kordestan, Kermanshah, Fars, Kerman and Sistan va Balochestan provinces was very weak. This article provides good information for transportation planners to recognize the critical provinces in terms of 14 relative indexes to enable the efficient allocation of limited resources.

## References


1. Evans, L., (2004). Traffic Safety. Science Serving Society, Mich Bloomfield.
2. World Health Organization (WHO), (2001). Burden of disease project. Global burden of disease estimates for 2001. <http://www3.who.int/whosis/ menu.cfm?path=burden>
3. Fancello, G., Carta, M., Fadda, P. (2014). A Decision Support System Based on Electre III for Safety Analysis in a Suburban Road Network. 17th Meeting of the EURO Working Group on Transportation, Sevilla, Spain, July 2-4, 2014, 175-184.
4. Khorasani, G., Mirmohammadi, F., Motamed,H., Fereidoon, M., Tatari, A., et al., (2013). Application of Multi Criteria Decision Making Tools in Road Safety Performance Indicators and Determine Appropriate Method with Average Concept. International Journal of Innovative Technology and Exploring Engineering , 3(5),173-177.
5. Lin, Y., Ye, Y., Niu, J., (2011). PCA-CA Method for Road Traffic Safety Macroscopic Evaluation and SafetyLevel Classification. 11th International Conference of Chinese Transportation Professionals (ICCTP), Nanjing, China, August, 14-17, 933-940.
6. Ahmadvand, A., Abtahy, Z., Bashiri, M., (2011). Considering Undesirable Variables in PCA-DEA Method: A case of Road Safety Evaluation in Iran. Journal of Industrial Engineering International, 7(15), 43-50.
7. Haghighat, F., (2011). Application of a Multi-Criteria Approach to Road Safety Evaluation in the Bushehr Province, Iran. Journal of Traffic and Transportation Research, 23(5), 341-352.
8. Wei, L., Sun, J., (2011). Evaluation Method of Road Traffic Safety Based on the Combined Principal Component Cluster Analysis. International Conference on Business Management and Electronic Information (BMEI), Guangzhou, May 13-15, 103-107.
9. Molla, M. M., Stone, M. L., Lee, E., (2015). Identification of Road Traffic Fatal Crashes Leading Factors Using Principal Components Analysis. Journal of URISA, North Dakota State University, Fargo, ND(Under Review).
10. James, G., Witten, D., Hastie, T., Tibshirani, R., (2013). An Introduction to Statistical Learning with Applications in R. 1th. Edition, New York.
11. Ertugrul, I., Karakasoglu, N., (2007). Performance evaluation of Turkish cement firms with fuzzy analytic hierarchy process and TOPSIS methods. Journal of Expert Systems with Applications, 36(1) , 702-715.
12. Mohammadi, A., Mohammadi, A., Aryaeefar, H., (2011). Introducing a New Method to Expand TOPSIS Decision Making Model to Fuzzy TOPSIS. Journal of Mathematics and Computer Science, 2(1), 150-159.
13. Iran Road Maintenance and Transportation Organization (RMTO), 2013. Annual Report <http://www. rmto.ir/NewTTO/DynaStat/DynaStatF>.


[ Downloaded from mail.ijaor.ir on 2026-11-27 ]