Available online **www.ijpras.com**

International Journal of Pharmaceutical Research & Allied Sciences, 2018, 7(3):63-73

Research Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Classification and Identification of IDP Camps After Mosul Events Based on Epidemics and Other Factors Using Cluster Analysis and Discriminant Analysis

Suhad ALi Shaheed AL-Temimi[1*], Rawaa Salh AL-Saffar[1], Hanaa Saad Shebib[2]

[1] Statistics Department, Collage of Economic & Administration, Mustanisiryah University, Iraq
[2] Faculty of Science, University of Kufa, Iraq
**E-mail:dr.suhadali @ uomustansiriyah.edu.iq*

ABSTRACT

Over the past decades, Iraq has witnessed an important population movement, and in many periods, it has been characterized by a high rate and expansion of its repercussions. This movement has included internal migration from rural to urban areas, and from small towns to big cities, it also included significant number of uprooted population inside Iraq, and the dense migration to outside of Iraq. The phenomenon of displacement is one of the most prominent population phenomena witnessed by Iraq in recent decades, which led to the masses of the population of a number of areas to move to other places and regions. The reason of this displacement of high incidence has been violence and armed attacks. The current humanitarian crisis did not happen suddenly. With the fall of Mosul on June 10th, 2014, a series of parallel and interlocking developments caused this large and unexpected event, including the ability of terrorist gangs to control up to one third of Iraqi territory, and the subsequent waves of displacement. In this research, multivariate statistical analysis methods were applied. The Cluster Analysis method was used to identify the disparities in the distribution of IDP camps among the Iraqi governorates after Mosul events in terms of the number of displaced families, and the distribution of families. A multivariate discriminating analysis was also used to identify the type of displacement camps (tents and caravans) that contribute significantly to this displacement variance.

Key words: *Displacement, Displaced Families, IDP Camps.*

INTRODUCTION

Over the past decades, Iraq has witnessed an important population movement, and in many periods, it has been characterized by a high rate and expansion of its repercussions. This movement included internal migration from rural to urban areas and from small cities to larger cities, and significantly the displacement of important numbers of the population inside Iraq, and the migration of the scouts out of Iraq. The phenomenon of displacement is one of the most prominent population phenomena witnessed by Iraq in recent decades, which led groups of residents of a number of areas move to other places and areas by increased cases of violence and armed attacks. The current humanitarian crisis has not suddenly happened. The fall of Mosul on 10th of June 2014 was a series of parallel and interrelated developments that resulted in this large and unexpected event [1]. The renewed conflict in Iraq has led to waves of displacement. In the years following the US invasion of Iraq in 2003, millions found themselves forced to leave their homes. Between February 15th, 2006 and April 2008, only 1.5 million people tried internal displacement. Hundreds of thousands fled to Jordan, the Syrian Arab Republic and other countries of the region. Between September 2013 and January 2015, more than 2.1 million people were displaced by the insurgency led by the organization of the Islamic state in Iraq and Syria, and renewed fighting throughout the country. This number has

been added to about 254,215 Iraqi refugees and 1.1 million internally displaced persons as a result of previous wars and political persecutions. Cases of forced displacement of the population were witnessed during the 1980s in the era of (Saddam Hussein). Of those affected by renewed violence in Iraq, more than 246,000 were Syrian refugees, and tens of thousands of Iraqi refugees have returned to their country in search of relative safety away from war in the Syrian Arab Republic. These moves indicated the multiple waves of displacement and the challenges faced by the actors in the field of humanitarian action who tried to develop plans for organizing displacement movements in the context of acute conflicts and insecurity. In this research, methods of multivariate statistical analysis were applied. The cluster analysis method was used to identify the disparities in the distribution of IDP camps among the Iraqi governorates after the Mosul events in terms of preparing displaced families for those governorates. The research found a close relationship between the governorates.

The problem of study

Diseases among the displaced are spreading in the camps because of the lack of health care and the overcrowding of the camps with large numbers of displaced people, especially under malnutrition and water scarcity. Nutritional research has provided clues on how foods or their supplements have the ability to optimize health and control diseases [2]. Diseases such as scabies and leishmaniosis have spread in the media, this problem has happened because of the absence of health care, food and water problems in the camps, and the lack of adequate sanitation or public facilities to maintain health. Scabies is one of the most important epidemic diseases in developing countries infecting males and females especially in rural areas [3]. And, generally, healthcare providers face many hindrances while maintaining the patients' and inpatients' safety [4]. This problem would be intensified in enviroments where standard facilities for healthcare are lacking or limited.

This study has presented important ways for the developments that took place in Iraq in the period after 15 January 2014, and the wave of mass exodus to Iraqi provinces. In that time, the fast provision of services to the displaced has become difficult because of the increasing number of displacement as a result of violent events and poor security in their areas. Therefore, the research focused on the classification of IDPs according to their numbers and areas of displacement through the use of multivariate analysis method, using the method of qualitative analysis to distinguish the type of displacement camps in terms of quality (tents or caravans), which contribute significantly to the disparity in the areas of displacement.

The objective of the study was to use statistical methods in multivariate analysis to classify and distinguish IDPs according to their numbers and the nature and type of camps, and the distribution of these families in the Iraqi provinces relying on the data adopted by the International Organization (International Organization Migration) [1].

The importance of the study

In the light of the events that passed through Iraq after the invasion, and the displacement of thousands of families from their areas moving to Iraqi provinces, it is important to identify the nature of the displacement of these families, the quality of the camps, and the provided inputs. All these factors affected the distribution of families and the variation in numbers from one governorate to another. The study has given a vision for the regulation of the governorates and the degree of convergence and disparity between them. The importance of the study is also evident in the need to highlight the problems in the distribution process and how to create a parallel according to the nature of each governorate. And, it is thus the first gesture in dealing with those important and fundamental problems that are exposed to Iraq.

The sample of the study

The data were based on reliable sources of government and international organizations within the European Union. These bodies have shown constructive and efficient assistance in reaching important results that can be disseminated and clarified within the active studies. The data were collected from the Ministry of Displacement and Migration, Information and Research Department, Statistics Department and the Information Bank of the camps in Iraq displaced by type of camps (tents, caravans) and in terms of number of displaced families. The other source was the International Organization for Migration (I.M.O), which had a prominent role in providing the latest data and transferring displaced families between the governorates between the invasion to Mosul, 15 January 2014 and 17 February 2017. The most recent migration tracking matrix of the International Organization for Migration (I.M.O) in Iraq identified 3,418,322 internally displaced Iraqi (569,772) families from January 1th, 2014 to March 31st, 2016), between March 2nd and 31st, 2016, an increase in displacement was recorded in Anbar (48,378) and Salah al-Din (23,718) due to military operations. For March 31st, 2016, the total number of IDPs reported was from eight Iraqi governorates of 18, most of them originally from Anbar (43 or 1,486886) and Nineveh (33%, or1,125,414).

The hypotheses of the study

Two hypotheses were formulated as follows :

1. There is no homogeneity in the distribution of IDP camps among Iraqi governorates.
2. The heterogeneity in the distribution of displaced persons is influenced by the type of camps (caravans, role of housing structures).

Theoretical side

CLUSTER ANALYSIS

Cluster analysis is one of the branches of multivariate statistical analysis, and a procedure to classify a group of situations or variables in specific ways [5-8], so that the cases classified within a particular cluster are homogeneous in relation to specific characteristics, and different from other cases found in another cluster.

Element X1 is a vector in the scale vacuum of n dimensions1.1

$$X_1=(X_1\ X_2\ldots\ldots\ldots X_{in}) \quad (1)$$

The elements are numerical values of possible measurement quantities.

Distance (D) is the space between the two elements, and the relationship between similarity and distance is an inverse relationship, and can be conducted by cluster analysis based on any of them. There are multiple formulas to measure the distance between them. The Minkowski Metric formula is as follows [9].

$$d(x,y)=[\sum_{i=1}^{p}|X_I - Y_I|^r]^{\frac{1}{r}} \quad (2)$$

When r 1 d (x, y) measures the distance of the urban centers between two points in a space with P of dimensions.

r 2 d (x, y) measures the distance between two points in a vacuum with P of the dimensions, and generally the change of value r leads to a change in weights giving many differences including small differences.

Cluster is a group of fairly homogeneous elements which describe what is inside the cluster, and are different from the elements within the other clusters [10].

Tree diagram of dendogram is the hierarchical form produced by the process of clustering (cluster formation) [10] and can be accessed in two ways : agglomerative, where analysis begins with one cluster per case. The clusters are then gradually aggregated until the specified number of clusters is reached.

Divisive : In this type, all cases are considered to be clustered in a single cluster and then the cases are classified into smaller and smaller clusters. In both cases, the results shown by the two methods is a hierarchical tree and the beginning of the division is called the root, and the so called nodes, and the final or final nodes on the tree. In trees with no branches, they are called the leaves and they represent the elements that met with each other. Each of the nodes in the tree, including the root, represents a qualitative set of objects that can be accessed in that node towards the foreground and through the tree.

Clustering steps

Calculate the distance matrix, correlation matrix, or similarity matrix.

1. The two elements with the shortest distances within the matrix calculated in the previous step shall be linked. In the case of equal distances, it is possible to connect to more than two elements at one stage (both elements).
2. The new distance matrix is calculated taking the changes in the second step into account.
3. Continue the binding process until access to the cluster tree. It is worth mentioning that the conversion of the data such as logarithm, or the conversion to the standard was carried out before the above operation, especially when there are different units of measurement.

Methods of cluster hierarchical

There are many methods of cluster analysis. There are certain characteristics that are different from the other methods. Some depend on the method of assembly, and the others depend on the method of fragmentation. The most common method of hierarchical analysis has been explained in this study. These methods do not require knowledge. The number of clusters are classified on the basis of the cases as they fit in relatively small samples. In cluster analysis, especially the so-called Linkage methods, these methods are suitable for the aggregation of cases as appropriate for the aggregation of variables, which is not true for other assembly methods.

Single linkage

It is also called the nearest neighbor. This method is mainly based on the fact that the two most similar elements form the nucleus of the cluster, and then the rest of the units are added to this nucleus sequentially. According to the degree of similarity with the elements of the nucleus of the cluster the most similar is added and then the least similar clusters are added gradually. And inking clusters with each other is based on the nearest distances or coefficients of symmetry between pairs of elements. According to the following formula, the rest of the units to this nucleus in sequence and the degree of similarity with the elements of the cluster nucleus are calculated :

$$d_{ij} = \text{Min}(d_i) \quad (3)$$

Where $_{ij}$ represent the elements in clusters I, j, respectively.

The Comprehensive Linkage method

It is also called the farthest neighbor. This method is mainly based on the fact that the two most similar elements between the elements form the nucleus of the cluster. This method is quite contrary to the principle of the work of the previous method. The elements candidate to enter the cluster should be in a max distance from any elements of the cluster. It is linked as follows :

$$d_{ij} = \text{Max}(d_i) \quad (4)$$

Where $_{ij}$ represent the elements in clusters I, j, respectively.

There is also the average linkage method and the incremental linkage method (ward`s).

Discriminated Analysis

Distinguished analysis is a statistical method for the analysis of multivariate data so that it deals with the issue of differentiation between two or more groups which are similar in many characteristics on the basis of several variables through the use of the distinctive function [5, 8, 11, 12]. It is a linear structure of independent variables. The qualitative analysis of the cluster analysis differs in that the idea of cluster analysis appears without prior knowledge of the number of groups or any of the vocabulary belonging to this group. Just as the qualitative analysis is different from the regression analysis. The dependent variable in the meta-analysis is a nominal variable, which is a qualitative variable, whereas the dependent variable in the regression analysis is often a constant variable, without the quantitative variables. The classification process is the subsequent process after the configuration of the distinctive function. This function is based on the prediction and classification of the new vocabulary of one of the groups under study with the least possible classification error. The data is equal to the groups studied. There is a case-by-case distinction of two groups. Non-linear discrimination is used in the case of unequal data [6].

The linear discriminator function

The function of discrimination is a model that can be formulated based on the kinds of indicators that have been selected and placed in two different groups. By this function, the singularity can be tested and the family will be limited to any group. If we assume that the area of the sample is W, it will be divided into two parts (R) from the first group, and W-R returns to the second group. The boundary between the two groups can return to any of these groups if there are variables of (x1 .x2 .... xp). The general formula of the distinctive function is as follows :

$$*\ldots\ldots* \text{ BpXp} \qquad {}_{r}=1.2 \quad Y_1=B_1X_1+B_2X_2 \quad (5)$$

Where P is the number of internal variables in function B, and the parameters of the standard distinguished function. To determine the differences between the two groups, it is appropriate to extract the computational circles of these two groups and the process estimation process (B.S), which makes the function give the best distinction between the two groups. In order to do this, the square difference between the two groups' average and the common variance of the two groups must be as large as possible.

$$\frac{[\mathcal{B}^r(\ \ell_1-\ell_2)]^2}{\mathcal{B}^1\mathcal{S}\mathcal{B}} Q = \frac{[\ \ell_1-\ell_2]^2}{\sum_{i=1}^{2}\sum_{j=1}^{ni}(y_{ij}-\ell_1)^2} \quad (6)$$

Where we estimate the characteristic function information by maximizing the Q ratio by partially dividing it by zero and obtaining

$$B\ ^{\wedge}=S1\ (\overline{X}_{1} - \overline{X}_{2})^{1} \quad (7)$$

In the case of two groups, we have only one distinguishing function. In the case of three groups, we have two distinct sets. After extracting the coefficients B ^, the class is classified into one of the groups based on the L-point, which makes the probability of the wrong classification as low as possible.

$$L = \frac{\ell_1 + \ell_2}{2} \quad (8)$$

The view is classified into the first group if ý> L is classified into the second group, and if ý<L is classified randomly to the first or second group if ý-L

$$Ý=(\overline{X}_{1} - \overline{X}_{2})S^{-1}X \quad (9)$$

Moral characteristic function test

When distinguishing between two groups, the hypothesis that the average of the two groups is equal can be tested. [9, 10]

$$H_0 : M_1\text{-}M_2$$
$$H_1 : M_1 \neq M_2$$

The test score used in the case of the distinction between two groups is T2 (Hoteling), and its formula is as follows :

$$T^2 = \frac{n1n2}{1n1+n2}D^2 \quad (10)$$

Where D2 represents (Mahalanobis Distance) and its formula is as follows :

$$D^2 = (\overline{X}_{1} - \overline{X}_{2})^1\ S^{-1}\ (\overline{X}_1 - \overline{X}_2) \quad (11)$$

The test (F) is used and its formula is as follows :

$$F = \frac{n1+n2-P-1}{(1n1+n2-2)P}T^2 \quad (12)$$

With a degree of freedom (P1n1-n2-p-1), H0 with a significant level of Fa was rejected for :

$$F_{cal} > Fa_{(P1n1\text{-}n2\text{-}p\text{-}1)}$$

H1 was accepted. This indicated that the mean of the groups was not equal, and that there were significant differences between the two groups. This means that the linear characteristic function was highly distinguishable. The WILKS-CRITERIA scale can also be used according to the following formula [9] :

$$\Lambda = \frac{|W|}{|T|} \quad (13)$$

Where T is the total variance and heterogeneity matrix of group,
And W is the variance and heterogeneity matrix within the group.
The value of ∧ between zero and one, if close or equal to one, indicates that the mean of the groups is equal, so there is no distinction between the groups. If the value is close to zero, it indicates the strength of the discrimination and the X ^ 2 scale can be used. This measure is more accurate than the ∧ scale and is as follows :

$$X^2 = -log(\Lambda) \quad (14)$$

(K-1) P where P is the number of variables, K is the number of groups.

Probability of classification error

There are two types of probability of misclassification :

1. The probability of error of classification : P12, which is the probability of classifying the individuals into the second group, is originally from the first group.
2. The possibility of error correction : It is possible to classify the individual into the first group, which is originally from the second group. The rating probability will thus be estimated as follows :

$$P_{12}=P_{21}-\Phi(-D/2) \quad (15)$$

Where
Φ) represents the normal distribution function, and
D is the root of the D2 (Mahalanobis Distance) scale.

The applied side

In this section, the cluster analysis method applied to the data after it has been classified according to the Iraqi governorates, which were displaced in order to find clusters of governorates that are homogeneous by classification and comparison at the same time. In addition, the method of discriminatory analysis will be applied to identify the types of camps (tents, caravans, structures, houses, hotels) that cause heterogeneity between the governorates to which the displaced are distributed. The results of the analysis and the comparison of the cases (governorates) have been shown below.

The study was based on variables (type of camping) as shown in Table (1). The Statistical Package for Social Sciences (SPSS ver.20) was used in all stages of the analysis.

Table 1. The distribution of displaced people in governorates by the type of camps

hotel	house	Caravan	camp	governorate
0	0	3404	16097	Erbil
0	0	0	2000	Mosul
0	0	0	39997	Dohuk
0	0	1520	3613	Sulaymaniyah
0	1012	36	6644	Kirkuk
0	0	39	31	Maysan
0	0	67	0	Dhi Qar
0	0	800	0	Wasit
0	0	0	17	Basra
0	0	1300	0	Karbala
396	1603	634	18570	Anbar
0	200	0	1410	Salahuddin
0	1791	1352	2797	Baghdad
0	0	1828	550	Diyala

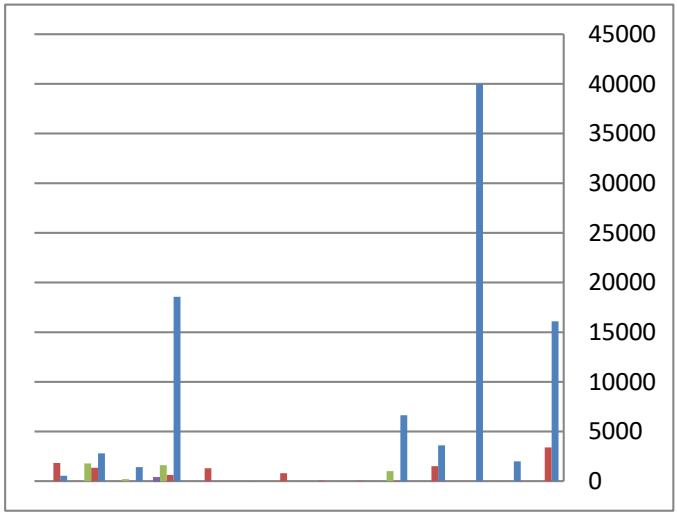

Figure 1. The distribution of IDPs according to the type of camps

The overall linking method was used to find the relationship between the governorates based on the similarity scale used in Minkowski Metric using the formula (2) when r=1. The results are shown in Table (1) represents the matrix

of Proximity Matrix determining the similarity or lack of similarity between the provinces, and expressing the distance derived between each province with the other provinces. The results of the cluster analysis were analyzed in order to verify the first hypothesis of the research which states that (there is no homogeneity in the distribution of IDP camps between the Iraqi governorates), the method of linking the comprehensive to find the relationship between the provinces was applied based on the similarity scale used by the Minkowski Metric coefficient by applying the formula (2), where (r =1). The results are shown in Table (2), which represents the Proximity Matrix, determining the similarity or lack of similarity between the governorates which is expressed by the derived distances between each governorate with the other governorates.

The provinces of the country were represented in the sequence of (1-14) according to the following sequence : Arbil 1, Nineveh 2, Dohuk 3, Sulaymaniyah 4, Kirkuk 5, Maysan 6, Dhi Qar 7, Wasit 8, Basra 9, Karbala 10, Anbar 11, Salah al-Din 12, Baghdad 13, and Diyala 14

As can be seen in Table (2), the governorates of Dhi Qar and Maysan formed the first concentration. The lowest coefficient was (0.001) in a new stage which was the next phase (2). It was concentrated in the governorates of Maysan and Basra. A new phase (3) was held in (Maysan and Ninewi) with a new approach coefficient of 0.033 to form a cluster in phase 4 with an approach coefficient of 0.111 to form an ascending cluster where the largest coefficient of approach was 22.498 between Anbar and Erbil governorates. The Dendorogram includes the measurement (levels) extending to 13 units of measurement, where the length of the line indicates the increase in degrees of similarity and there are several nodes (Nodes in the tree, each node of the merger representing two or more as shown in Figure 2). At the end of the data analysis, the results were placed in clusters. In terms of the similarities in the sources, the types of camps in the provinces are shown in Figure (2). There are governorates which are at close distances (Maysan, Dhi Qar, Basrah, Nineveh, Salahuddin, Karbala, Sulaymaniyah, Diyala, Wasit). This reflected the similarities in the type of camps in each governorate, and the large homogeneity between these governorates, thus constituting the first cluster with the least number of displaced people. While five provinces are held at different distances (Kirkuk, Baghdad, Irbil, Dohuk and Anbar).

Table 2. The aggregation of the governorates according to the stages of cluster analysis

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	(Maysan)6	Thi qar(7)	0.001	0	0	2
2	(Maysan)6	Basra (9)	0.003	1	0	3
3	(Ninevah)2	Maysan (6)	0.033	0	2	4
4	(Ninevah)2	Salahuddin (12)	0.111	3	0	7
5	(Sulaymaniyah)4	Karbala (10)	0.149	0	0	6
6	(Sulaymaniyah)4	Diyala (14)	0.222	5	0	8
7	(Ninevah)2	Wasit (8)	0.627	4	0	8
8	Ninevah 2	Sulaymaniyah (4)	2.102	7	6	10
9	(Kirkuk)5	Baghdad 13	3.302	0	0	10
10	(Ninevah)2	Kirkuk 5	6.111	8	9	11
11	(Erbil)1	Ninevah 2	10.685	0	10	12
12	(Erbil)1	Dahuk(3)	13.547	11	0	13
13	(Erbil)1	Anbar(11)	22.498	12	0	0

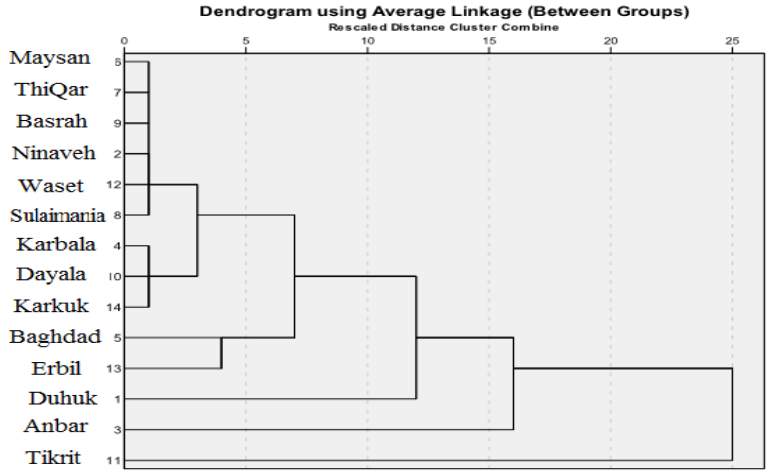

Figure 2. The tree diagram of the method of linking the comprehensive link between the provinces

The differential analysis of the results

The differential analysis was based on the composition of the linear discrimination function on several variables. The number of displaced people in each type of camp was collected for each individual governorate. The totals obtained from the cluster analysis results were used in the discriminatory analysis. Since the number of totals was two, the differential analysis was used in case of two dependent variable groups, which represented the state of the governorate (1 represented the governorates of the first cluster with the lowest number of displaced by the type of camps, and 2 represented the provinces of the second cluster of high income). Additionally, the explanatory variables were the types of camps within each province, which are distributed to the displaced people (X1 tents), (X2 caravan), (X3 structures and houses), (X4 hotels).

Testing the significance of the variables in the characteristic function

It was needed to calculate the arithmetic mean of variables in groups, as shown in Table (3).

Table 3. Arithmetic averages and standard deviation for the preparation of displaced people by the type of camping for the two groups and the total.

yy		Mean	Std. Deviation
1.00	camp	7599.8889	13229.92051
	caravan	651.7778	1157.66973
	house	112.4444	337.33333
	hotel	.0000	.00000
2.00	tent	4665.4000	7844.24826
	caravan	1022.8000	712.53856
	house	718.8000	899.15555
	hotel	79.2000	177.09658
Total	tent	6551.8571	11347.83691
	caravan	784.2857	1007.46860
	house	329.0000	640.07572
	hotel	28.2857	105.83545

Table (3) illustrates the significance of all variables determining the importance of each variable in the distinguishing function, and table 4 shows their effects on the analysis of the results using the single-variance analysis.

Table 4. The test (F) for each variable in the linear characteristic function

Variables	Wilks' Lambda	F	df1	df2	Sig.
tent	0.488	28.319	1	12	0.011
caravan	0.498	31.632	1	12	0.025
house	0.862	5.422	1	12	0.089
hotel	0.776	3.929	1	12	0.190

The table above shows that all variables (X1 tents and X2 caravans) are highly significant and have a significant effect in terms of distinguishing between the two governorates. The variables (X3 structures and roles) and (X4 hotels) have no significant effect.

Moral function significance test

In this table, the difference between the two groups for all the characteristic variables was explained by a discriminant function. The discriminant function was interpreted as 100% of the variance, and its legal correlation coefficient reached to (0.855). In (13), the value of this measure was close to zero, also in the scale of 2 according to the formula (14), the significance of this measure was observed. Therefore, it can be said that the discriminatory function was sufficient to explain the discrepancy between the two groups. This proved the possibility of a discriminatory equation for depicting an optimal linear model consisting of a set of variables (number of displaced by camp type), capable of distinguishing between the two groups.

Table 5. Benchmarks used in the function test

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
	0.505[a]	100.0	100.0	0.855
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
	0.092	22.256	4	0.023

Explanation of Standard Distinguished Transactions Standardized parameter

The coefficients were estimated in accordance with the formula (7), their values are shown in Table (6). The standard differential coefficients were of great analytical significance, where the standard differential coefficient of the variable expresses the amount of its contribution. The variable in the discriminatory equation was significant if the absolute value of its coefficient was large, and the standard differential coefficient signal indicated the variable as either negative or positive.

Table 6. The parameters of the standard characteristic function

variables	coefficients
tent	3.540
caravan	2.435
house	-0.675
hotel	-0.496

Table (6) shows the discriminatory weights distinguishing the variables in the discriminant function. Note that the variable (X1 tents) had a high effect compared to the other variables, and a positive contribution in distinguishing the provinces followed by variable (X2 caravans). A positive contribution was also made to the variables (X3 structures and houses) and (X4 hotels).

The probability of correct classification

The classification of the governorates by using a discriminatory analysis was checked to find whether they actually fall within the suction group or the province belongs to the other group. The result showed that the classification of the province was wrong. The average of the two groups is shown in Table (7).

Table 7. The results of the classification

y			Predicted Group Membership		Total
			1.00	2.00	
Original	Count	1.00	10	0	10
		2.00	0	4	4
	%	1.00	100.0	.0	100.0
		2.00	70.0	30.0	100.0
A 98.6% of original grouped cases correctly classified.					

It is clear from the table that the correct classification rate for all groups was (98.6), ie 14 out of 14 governorates were classified correctly into the two groups to which they belong.

CONCLUSIONS

1. The results of the cluster analysis showed that Iraqi governorates that were displaced were classified into homogeneous groups according to the camp types, thus, classified into homogeneous clusters.

2. There was a convergence in the types of camps for (9) Governorates formed in the first cluster.
3. While there was a convergence of camp types between five governorates, which were located at different distances (Kirkuk, Baghdad, Irbil, Dohuk and Anbar), they included the second largest number of displaced people.
4. Discriminatory analysis was used to distinguish individual income sources. The aggregates extracted from the cluster analysis results were used in the discriminatory analysis.
5. The variable (X1 tents) had a high impact compared to the other variables and a positive contribution to the distinction of the provinces followed by the variable. It was a positive contribution to the distinction of governorates followed by variable (X2). A positive contribution was also made to the variables (X3 structures and houses) and (X4 hotels).
6. The correct classification rate for all groups was (98.6) i.e, (14) governorates out of (14) governorates have been properly classified into the two groups that they belong to.

Recommendations

1. The variable (X1 tents) and the variable (x2 caravans) are the sources of the basic difference, which contributed greatly to the distinction between the governorates. The researchers recommend that the leaves should be reduced between the provinces as the camps and caravans are more stable for the displaced people. Because they are in the form of gatherings, and there are services provided by local and foreign bodies, as for the role and hotels, they need fixed amounts and income so that the displaced can live and settle in, which is not applicable to most of the displaced.
2. Cluster analysis should be used in the classification of governorates and the country in general, according to the statistics of the International Organization for Migration (IOM).
3. Discriminatory analysis should be used in the classification of many phenomena, whether economic, social or health of the category of displaced and forcibly displaced.
4. This study was conducted in the case of the implementation of a new survey for the displaced after the operations of Mosul fighting. Other studies can be done in other contexts

REFERENCES

1. International Organization for migration versions (2014-2017) (IOM).
2. Jalaja, K. D. & Santhi Sri, K. (2013). Health Promoting Properties of Phytochemicals. International Journal of Pharmaceutical and Phytopharmacological Research, 2(5), pp. 307-311.
3. Mohy, A. Jaloob Aljanaby, A. A. Al-Hadraawy, S.K. (2018). Serum concentrations of CD4+ and CD8+ in patients infected with scabies caused by Sarcoptes scabiei. Journal of Advanced Pharmacy Education & Research, 8(1), pp.121-126.
4. Al-mazroea, A.H. AlturkiA. M. (2017). Frequency of medical errors complaints against medical health centers in Al Madinah Al Munawarah region of Kingdom of Saudi Arabia. International Journal of Pharmaceutical Research & Allied Sciences, 6(2), pp. 308-312.
5. Fareed and Ghanem, Adnan, 'Multivariate Statistical Analysis in the Characterization and Distribution of Families within the Socio-Economic Structure of Society,' Damascus University Journal of Economic and Legal Sciences, Volume 23, No. 2, 331), 2007.
6. Al-Jabouri, shalal and Hamza Saleh ' multivariate analysis ', books for the University of Baghdad, Baghdad, Iraq, 2000.
7. Khalid Khwaja, Zuhdi ' methods of data analysis aalsrh income and expenses ' Arab Institute for training and research c, Baghdad, Iraq, 2005.
8. Mustafa, Nizar ' using some methods of cluster analysis in classification with practical application, technical journal, technical college employees, Baghdad, vol (20). 2007.
9. Rencher, A. C "Methods of Multivariate Analysis", Second Edition, John Wiley & sons, New York, USA,2002.
10. Timm, N.H "Applied Multivariate Analysis" Springer-Verlag, New York, Inc. USA,2002.
11. Fazzo, L. et al "Cluster analysis of mortality and malformations in the Provinces of Naples and Caserta (Campania Region)", Ann Ist Super Sanità, Rome, Italy, Vol. (44), No. (1), PP (99-111),2008.

12. Gwo, F. L &, Chun, M. W "Performing cluster analysis and discrimination analysis of hydrological factors in one step", Advances in Water Resources, Department of Civil.