RESEARCH ARTICLE | DECEMBER 04 2023

The use of hierarchical clustering and the Markovian model in study tourism in Iraq

Muna T. Ghafil ✉; Wafaa A. Ashour

AIP Conf. Proc. 2834, 080019 (2023)
https://doi.org/10.1063/5.0161478

The Use of Hierarchical Clustering and the Markovian Model in Study Tourism in Iraq

Muna T. Ghafil [a] and Wafaa A. Ashour [b]

College of Administration and Economics, Statistics Department, Basra University, Basra, Iraq

[a] Corresponding author: muna.ghafil@uobasrah.edu.iq
[b] wafaa.ashoor@uobasrah.edu.iq

Abstract. In this research, the hierarchical clustering method has been applied to classify the Iraqi provinces in the tourism aspect, which depends on the multivariate analysis specified depending on the points of similarity and difference between the data, using the statistical program (spss.v.23), in addition to using the Markovian model to estimate the matrix of transitional probabilities in a non-educational way. It was found that the Markovian model is compatible with the cluster analysis and the ease of using the outputs of the cluster analysis in calculating the transitional probabilities matrix.
Keywords: Hierarchical clustering, Markov model, multivariate analysis, transitional probability matrix.

INTRODUCTION

Cluster analysis is defined as a set of mathematical methods used to determine the structural properties of statistical data, by classifying them into groups (within clusters), so that the vocabulary within each group is similar to each other for the approved characteristics, (so that the groups are apparently different from each other) In other words, the objective of the cluster analysis is to collect the sample items and classify them into groups that are internally homogeneous and externally differentiated between each other.

Classification has been used since ancient times for the purpose of classifying humans based on the characteristics of color, gender, and others. One of the most important research in which classification scales were used is in the year 1960, a measure of similarity was found by researchers (Curtis and Bary) and it was applied in the field of education studies and in In 1961, the researcher (Solcal) was able to find a scale called (distance Euclidian), which measures the distance between the two comparative elements. (Mustafa, 2007)

The theory of Markov processes occupies a large place in the theory of stochastic processes. This position enhances the multiplicity of applications that use Markov processes in physical and biological models, sociology, engineering and management science, in addition to their multiple applications in many statistical and engineering models. (Idrissi, 2013).

The Russian mathematician (Andreevich Markov 1856 - 1922) came up with a new idea to facilitate prediction calculations in the style of time series after he was involved in the prediction calculations all the observation values of the studied phenomenon (the random variable understudy), in addition to that, obtaining the predictive values is very difficult. When the series is long and with large numbers, and therefore Markov came up with a new method, which states [In the analysis of time series it is sufficient to rely on current values to predict the values of the series in the future and not to rely on previous values or historical values of the studied series. (Al-Shamrty, 2017)

RESEARCH PROBLEM

Tourism is of great importance in supporting the national economy because of the financial resources it achieves, which contribute to the development process, and also contribute to increasing job opportunities due to its ability to absorb large numbers of manpower in various tourism activities represented in hotels, restaurants, commercial markets and popular industries. Since the hotel activity and tourist accommodation is one of the most important requirements for work in tourism, it is necessary to know the disparity and difference in hotel services between the governorates and to know which governorates are close and which are different. To raise the level of services provided and thus develop the tourism aspect in the country.

2nd International Conference of Mathematics, Applied Sciences, Information and Communication Technology
AIP Conf. Proc. 2834, 080019-1–080019-10; https://doi.org/10.1063/5.0161478
Published by AIP Publishing. 978-0-7354-4715-8/$30.00

RESEARCH OBJECTIVE

The research aims to use the cluster analysis method in dividing the tourism data for the year 2018 in Iraq, into groups and each group has a great convergence between its vocabulary, and the link between the cluster analysis and the Markov model to estimate the transitional probability matrix of Markov. Tourism also has an important role in the progress of countries because of its importance in generating profits for the country, especially if the number of Arab and foreign guests increases, which causes foreign currency to enter the country.

RESEARCH IMPORTENCE

To research the importance of using multivariate statistical methods in the analysis that facilitates the process of comparison in tourism services between governorates by classifying them and knowing the extent of their homogeneity in the most important indicators of tourism of great importance in the progress of tourism, which helps in identifying the obstacles and imbalance that makes tourist hotels provide services with high efficiency, including Fits the requirements of the guests.

RESEARCH LIMITS

The research relied on the data available in the complete statistical set for the year 2018-2019 issued by the Central Bureau of Statistics.

PREVIOUS STUDIES

A study (Mustafa, 2007) used cluster analysis on the data of a group of Arab Maghreb countries in the African continent to reach the totals of countries that have a number of common characteristics to link the economic integration between these countries.

A study (Al-Shakurji, 2008) worked on the use of cluster analysis, hierarchical clustering method and segmentation method on 48 players in the College of Physical Education / University of Mosul.

A study (Habal, 2012) In this research, cluster analysis was used to study the chemical, biological and manufacturing properties of high-quality wheat flour and the high degree of similarity between samples to which oatmeal was added in properties.

Study (Al-Khayat, 2013) This research dealt with the use of the Markovian model in the issue of link prediction in the analysis of Internet networks, and a practical experiment was conducted to highlight the importance of compressing the transition matrix.

A study (Al-Halawani, 2018) where cluster analysis was used to classify the Egyptian governorates (27 governorates) for some of the characteristics that characterize these governorates.

Study (Al-Obaidi, 2018) The researchers used the time series of the number of pneumonia infections as a Markov series, and the stable distribution of this series was found.

THEORETICAL ASPECT (the statistical methods used)
CLUSTER ANALYSIS

It is a statistical method that includes many procedures for classifying a group of cases or variables in many ways and arranging them into clusters so that the cases classified within a particular cluster are homogeneous among themselves with specific characteristics and differ from other clusters.

Cluster analysis is divided into two main types: hierarchical cluster analysis and non-hierarchical cluster analysis.

HIRRARCHICAL METHODS

Hierarchical clustering methods start by assuming that each item belongs to a single cluster, and in each plan, the closest pair of clusters are combined until only one cluster contains all the items. There are many ways that depend on the process of collecting vocabulary or elements on the similarity matrix, including:

a- Single link method: It is also called the nearest neighbor method and depends on determining the distance between clusters with the smallest distance between any two items (nearest neighbor) in different clusters.

The distance is calculated as follows:

$$d\left(S_i, S_j\right) = \min\left\|x_i - x_j\right\| \quad (1)$$

Whereas $S_i = \{x_1, x_2, \cdots, x_n\}$ represents the first cluster, and $S_j = \{x_1, x_2, \cdots, x_m\}$ represents the second cluster.

b- Global linking method: It is also called the farthest neighborhood method, in which the distances between clusters are calculated with the largest distance between any two items (two elements) within the different clusters. It is also called the full link method because it starts with all the elements as single clusters and then the strongest connections between the vocabularies are added. The perfect link formula is calculated as:

$$d(S_i, S_j) = \max\left\|x_i - x_j\right\| \tag{2}$$

Whereas $S_i = \{x_1, x_2, \cdots, x_n\}$ represents the first cluster, and $S_j = \{x_1, x_2, \cdots, x_m\}$ represents the second cluster. (Gavira,2021)

C- The linkage pair rate method or the group pair rate: in which the distance between two clusters is calculated using the average pair wise distance between all pairs of items in the different clusters. It is an intermediate method between the two methods of linking Max and Min, and is represented by the following formula:

$$\Pr oximity(S_1, S_2) = \frac{\sum proximity(x_1, x_2)}{Size(S_1) * Size(S_2)}, \quad x_1 \in S_1, x_2 \in S_2 \tag{3}$$

Where S1 and S2 represent the first and second clusters.

D- The central clustering method: in which the distance between two clusters is calculated based on the centers of the clusters. There are two methods used for the central clustering method:

- Unweighted Group Pair Mean Center Method: In this method, the distance between two clusters is determined as the difference between two centers as in the following formula

$$d_{mean}(S_i, S_j) = \left|m_i - m_j\right| \tag{4}$$

$m_i = \frac{1}{n}\sum x_i$, $m_j = \frac{1}{m}\sum x_j$ They are the averages of the two clusters Si and Sj, respectively.

When the median is used instead of the center of the cluster,

$$d_{med} = \left|med_i - med_j\right| \tag{5}$$

$\left(med_i, med_j\right)$ They are the medians of the two clusters Si and Sj, respectively.

- Weighted group pair means center method: This method is similar to the previous method, except that there are weights that are taken into consideration when the number of elements in it varies.

e- Word method: it is sometimes called the least variance method because it uses the analysis of variance method to calculate the distance between clusters. This method is better than the rest of the previous methods. It is given by the formula:

$$d_{ward}(S_i, S_j) = n.m.d_{ij}^2 /(n+m) \tag{6}$$

d_{ij}^2 It represents the distance between cluster i and cluster j defined in the central clustering method. m,n are the number of elements in the two clusters in a row. In this method, the sum of the squares of error is calculated between each pair of items (vocabularies) and then the pairs that have the smallest sum of squares error (SSE) are connected, which is calculated by the formula:

n total vocabulary, n_j The number of vocabulary in the cluster j_{th}. (Hanish, 2019)

$$SSE = \sum_{j=1}^{n}\left[\sum_{i=1}^{n_j} x_{ij}^2 - \frac{1}{n_j}\left(\sum_{i=1}^{n_j} x_{ij}\right)^2\right] \tag{7}$$

Some concepts of cluster analysis:

Space: is the space or space that separates two words.

Cluster: A group of vocabulary that is homogeneous with each other and is different from the vocabulary in other clusters.

Relationship Matrix: It is a symmetric matrix whose number of rows is equal to the number of its columns. Where the cluster analysis begins with the formation of that matrix, whose elements represent one of the measures of the distance between observations, and the general form of this matrix is as follows:

$$D=\begin{bmatrix} d_{11} & d_{12} & \cdots & \cdots & \cdots & d_{1n} \\ \vdots & d_{22} & & & & \vdots \\ & \vdots & \ddots & & & \vdots \\ & \vdots & & \ddots & & \vdots \\ & \vdots & & & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & \cdots & \cdots & d_{nn} \end{bmatrix} \quad (8)$$

Where D is an array of degree $n \times p$. (Jawad, 2009)

Tree: It is the hierarchical shape resulting from the clustering procedure, which can be accessed in two ways:

Divisional Cluster Analysis: All variables are grouped into one cluster and then classified into smaller and smaller clusters.

Cluster aggregation analysis: the analysis begins with one cluster for each case, then the similar clusters are gradually grouped until we reach the required number of clusters.

Classification: It is the arrangement of things based on the similarity or difference between them. (Kadhim, 2009)

STOCHASTIC PROCESSES

It is a statistical method in which a stochastic process is defined as a set of random variables that changes with a change of given evidence, for example, time, whose theoretical values cannot be found with complete accuracy, but has a certain range of possible values that depends on a probability distribution that describes Its value for each value of the index. It is also known as a set of random variables $\{x_j, n \geq 0\}$ In terms of time (or any other evidence); it can also be defined as a series of random variables generated using probabilistic laws. Denotes random process $\{x_n\}$ where $\{n\}$ represents the discrete time $\{n = 0,1,\cdots\}$ it is symbolized by $\{x_n\}$ this is for continuous time $\{t \geq 0\}$. Hence, the process $\{x_n\}$ is called the stochastic process the values assumed by it are called states, all possible values are called state space, and all possible values of the index parameter are called parameter space, which can be discrete or continuous, where The parameter T usually represents time, for example, the number of traffic accidents, which is a random variable, but when we measure it in terms of time or road length, then we have a random process represented by the number of time accidents or the length of the road distance. (Al-Rubaie, 2005)

MARKOV MODEL

A stochastic process with a discrete parameter $\{x_t : t = 0,+1,+2\cdots\}$ or with a continuous parameter $\{x_t : t > 0\}$ It is called a Markov process if it has the following property:

$$P_{jk}=P_r\ \{x_{n+1}=k\ /x_n=j,\ xn_{-1}=j_1,\ xn_{-2}=j_2,\ xn_{-3}=j_3, \ldots,\ x_0=j_{n-1}\} \quad (9)$$

since $x_{n-1}=j_1$, $x_{n-2}=j_2$, $x_{n-3}=j_3$, ... , $x_0=j_{n-1}$ neglected according to Markov logic, that is, in equation 9 can be written according to Markov's logic, as follows:

$$P_{jk}=P_r\ \{x_{n+1}=k\ /\ x_n=j\}. \quad (10)$$

To study any phenomenon or system, we go to know the state of the system at time t1, for example, where it is independent of all other states before time t0 and depends on the previous states of the system, i.e. at time t0, so the random process that achieves this condition is called Markov.

In more detail, it means the conditional probability of the random process of the variable xn+1 given all its past values, which in this case depends on xn only and does not depend on any value in the past and this is the Markov property that depends on causal probability. The classification of Markov processes according to the state of the system is either discrete or continuous, as is the case for parameter space. (Tag 2007)

MARKOV CHAINS

A Markovian process that takes a discrete state space is called a Markov chain, excluding the parameter space state as discrete or continuous and uses the set of integers [S=0,1,2,3,4,....] to represent the state space as the Markovian chain is known as A time-discrete stochastic process in which a random variable xn is related to the variable before it and affected by the change after it. It is also expressed as a Markov chain as a chain consisting of random variables to which the aforementioned Markovian property applies.

TRANSITION PROBABILITIES

Markov chains are built on the basis of the transition of the phenomenon from one state to another within a state space controlled by probabilistic laws, and the probabilities that describe the transitions between states of the Markov chain during a specified period of time are called transitional probabilities. The transition probability (going one step) from state j or (value) in time (n) to state or value (k) at time (n+1) is denoted by p_{jk} assuming that these probabilities are stationary over time. This applies to the discrete time Markov chain, but if the parameter space is continuous, the transition between states becomes through probability densities instead of transition probabilities. The transition probabilities are placed in a square matrix where two conditions must be met:

1. $\sum_{k\in E} p_{jk} = 1$ for all values of j of E.

2. $P_{ij} \geq 0$ for all values of j, k.

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ . \\ . \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & . & . \end{array} \\ \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{05} & . & . \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & . & . \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & . & . \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & . & . \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & . & . \\ p_{50} & p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{bmatrix} \end{array}$$

For example (P01): It represents the probability that the process is in state zero and will move to state 1 and thus the rest of the probabilities are explained. (Nicolas, 2018)

STATIONARY AND STEADY STATE

Stationary and steady state is generally means that the statistical characteristics of the stochastic process do not change to one degree or another over time.

As for the state of stability, it appears when the stochastic process continues for a long time, as the ratios of the number of transitions for each state are stable at a certain value, and it is called the stable probabilities of that state, i.e. the behavior of $P^{m}_{(j,k)}$ appears.

When m→∞ then the stable distribution is defined as:

Probability Vector $u = [u_1^s, u_2^s, \dots, u_n^s]$ who achieves $u^s * P^m = u^s$, a stable distribution is called a stationary distribution when P is the matrix of transitional probabilities of Markov chains with m finite states, then:

$$\lim_{m\to\infty} P^m = U = \begin{pmatrix} u \\ u \\ \vdots \\ u \end{pmatrix}.$$

$$u = (u_1 \quad u_2 \quad \dots \quad u_n)$$

represents the only constant probability vector.

$$\sum_{j=1}^{m} u = 1 \ , \quad 0 \le u \le 1$$

The stable distribution for the coming period can also be found by the following formula: UP=U

That is, if m approaches ∞, the transition probabilities of m steps of the matrix P_{jk} will depend on the last state and not on the initial state, and this means that after the passage of a large number of attempts, the chain will reach the state of stability. (Al-Adhari, 1991)

PRACTICAL APPLICATION OF RESEARCH

The multivariate statistical method (hierarchical cluster analysis), which works to classify the Iraqi provinces into groups whose vocabulary has common characteristics, has been applied. For the purpose of knowing the work of each item and its correlation with other items in the same group (cluster), using the statistical program (SPSS v23).

The research sample included (15) governorates that represent the Iraqi governorates, except for the Kurdistan region, due to the lack of data for the region. The following indicators have been adopted

TABLE 1. It represents the indicators that were adopted in the research

Pointer	code	Pointer	code
Number of first-class hotels	X_8	Total tourist facilities	X_1
Number of first-class hotels	X_9	number of guests	X_2
Number of second-class hotels	X_{10}	number of employees	X_3
Number of third-class hotels	X_{11}	Total wages (thousand dinars)	X_4
Number of hotels of the fourth class	X_{12}	Revenue (thousand dinars)	X_5
Number of fifth-class hotels	X_{13}	Expenses (thousand dinars)	X_6
		Purchases (thousand dinars)	X_7

ANALYZING AND DISCUSSING CLUSTER ANALYSIS RESULTS

The hierarchical cluster analysis was applied considering that the data matrix represents the governorates and their columns, as they are the study variables installed in the above table.

First, the kinship matrix: to obtain the kinship matrix, the aggregation method based on Euclidean distance was used.

The kinship matrix is a symmetric matrix, as the elements above the main diameter are equal to the elements located below the main diameter. It is noted that the most degree of convergence was between Diyala and Salah al-Din, as the distance between them reached (0.009). As for the largest degree of separation, it was between Karbala and Anbar governorates, as the distance between them was (3.082) as shown in Table No. (2) On the last page of the search.

TABLE 2. Represents the matrix of kinship

	Euclidean Distance														
Case	1:ninwa	2:kerkuk	3:diala	4:anbar	5:baghdad	6:babilon	7:kerbala	8:wasit	9:saladyn	10:najaf	11:qadisya	12:muthana	13:thy-qar	14:misan	15:basrah
1:ninwa	.000	.204	.314	.314	2.762	.312	3.001	.237	.314	1.566	.239	.238	.312	.313	.587
2:kerkuk	.204	.000	.209	.212	2.762	.200	2.961	.136	.205	1.545	.148	.158	.192	.192	.576
3:diala	.314	.209	.000	.009	2.896	.026	3.081	.084	.009	1.690	.083	.079	.038	.031	.657
4:anbar	.314	.212	.009	.000	2.896	.025	3.082	.085	.010	1.690	.084	.080	.044	.033	.656
5:baghdad	2.762	2.762	2.896	2.896	.000	2.881	2.044	2.847	2.892	2.010	2.850	2.862	2.869	2.878	2.316
6:babilon	.312	.200	.026	.025	2.881	.000	3.059	.080	.021	1.669	.083	.081	.032	.019	.644
7:kerbala	3.001	2.961	3.081	3.082	2.044	3.059	.000	3.046	3.078	1.551	3.054	3.057	3.049	3.057	2.729
8:wasit	.237	.136	.084	.085	2.847	.080	3.046	.000	.081	1.642	.029	.034	.082	.079	.618
9:saladyn	.314	.205	.009	.010	2.892	.021	3.078	.081	.000	1.686	.082	.080	.038	.025	.654
10:najaf	1.566	1.545	1.690	1.690	2.010	1.669	1.551	1.642	1.686	.000	1.649	1.655	1.664	1.665	1.405
11:qadisya	.239	.148	.083	.084	2.850	.083	3.054	.029	.082	1.649	.000	.033	.085	.082	.619
12:muthana	.238	.158	.079	.080	2.862	.081	3.057	.034	.080	1.655	.033	.000	.081	.084	.628
13:thy-qar	.312	.192	.038	.044	2.869	.032	3.049	.082	.038	1.664	.085	.081	.000	.032	.639
14:misan	.313	.192	.031	.033	2.878	.019	3.057	.079	.025	1.665	.082	.084	.032	.000	.644
15:basrah	.587	.576	.657	.656	2.316	.644	2.729	.618	.654	1.405	.619	.628	.639	.644	.000

TABLE 3. Represents the distribution of vocabulary as members of clusters

Case	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1:ninwa	1	1	1	1
2:kerkuk	1	1	1	1
3:diala	1	1	1	1
4:anbar	1	1	1	1
5:baghdad	2	2	2	2
6:babilon	1	1	1	1
7:kerbala	3	3	3	2
8:wasit	1	1	1	1
9:saladyn	1	1	1	1
10:najaf	4	4	3	2
11:qadisya	1	1	1	1
12:muthana	1	1	1	1
13:thy-qar	1	1	1	1
14:misan	1	1	1	1
15:basrah	5	1	1	1

We notice that the first knot was between the governorates of Diyala and Salah al-Din, then a new knot was added, represented by Anbar Governorate. And so on, until the first group ended in Basra Governorate. As for the second cluster, which started with the first knot between Karbala and Najaf, another knot was added, represented by the province of Baghdad.

According to the hierarchical cluster classification in Table (3), and by applying the Markov model for the purpose of building a matrix of transitions in the dimensions 4 * 4 to represent the transitions between the states of the process by first building a Markov model by finding the number of iterations for the transition between the four classified clusters by calculating the number of transitions from state j to state k in one step. We get the following matrix of frequencies: (Al-Qurashi, 2013)

$$F = \begin{pmatrix} 11 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

So we have the following transition matrix:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

We note from the matrix P that each of the Al Najaf Al ashraf and Karbala governorates reaches a state of sufficiency in terms of interest in religious tourism for both governorates, as well as for the other governorates represented by Nineveh, Kirkuk, Muthanna, Qadisiyah, Wasit, Dhi Qar, Maysan, Babylon, Anbar, Salah al-Din, Diyala, they are sufficient governorates in terms of interest in the tourism aspect in them according to the movement and development of tourism there. As for the governorate of Basra, followed by the governorate of Baghdad, it enjoys a movement of advancement and development in the tourism reality, at a rate of 100%, to move to a state of sufficiency.

To calculate the long-term probabilities for each classification, we get the following probability vector:

$$U = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$$

We note from the probability vector U and according to the analysis of the data subject of the research hierarchically clustered and the application of the Markovian model on it, it becomes clear to us that Baghdad is the only Iraqi governorate in which work continues in the long term to develop the tourism sector, especially since it is the capital.

CONCLUSIONS

The researchers reached the following conclusions:

1- The governorates closest to the tourist side are Diyala and Salah al-Din. The most remote provinces are Karbala and Anbar.

2- Baghdad governorate has its own cluster because it is the capital and its tourism season continues.

3- Karbala and Najaf governorates form a knot of their own due to the presence of religious shrines in them.

4- We note that the Markovian model is compatible with the hierarchical cluster analysis from the analytical side and from the realistic side.

5- Ease of using the outputs of the cluster analysis in calculating the transitional probabilities of the Markovian model.

RECOMMENDATIONS

According to the previous conclusions, the researchers would like to put forward some recommendations:

1- Increasing the government's interest in the tourism aspect in the governorates (Nineveh, Anbar, Kirkuk, Babil, Salah al-Din, Maysan, Wasit and Muthanna).

2- The necessity of increasing researchers' interest in using statistical methods to study the tourism aspect in Iraq and comparing the results with the results of this study.

3- The researchers recommend the necessity of providing and updating data and indicators related to the tourism aspect to be accessible to researchers.

4- Increasing studies related to the tourism aspect because of the importance of tourism in the development of the country.

REFERENCES

1. Al-Adari, Faris Muslim, Al-Wakeel, Ali Abdul-Hussein, “Chained Operations”University of Baghdad, Mosul University Press, 1991 AD.
2. Al-Halawani ‘Maji Ahmed’ “Using the cluster analysis method in classifying the governorates of the Arab Republic of Egypt” The Arab Journal of Economics and Trade ‘Ain Shams University’ Vol. 48 ‘No. 1’ pp. 517-469 ‘2018.
3. Al-Khayat, Basil Younes and Hussein, Talal Fadel, “Using the Markovian model in predicting the connection with the application of the Iraqi Journal of Statistical Sciences, issue 23, pp. 33-50, 2013.

4. Al-Qurashi, Ihsan Kazem, Abdullah, Majid Jaseb, "Nonparametric estimation of the transitional probability matrix using cluster analysis", Maysan Journal of Academic Studies, Vol. (12), Issue (22), June 2013.
5. Al-Obaidi ‘Abdul Ghafour and Suleiman ‘Ammar Yassin’ “Using Markov Chains in the Medical Fields” ‘The Iraqi Journal of Statistical Sciences’ Issue 27’ pp. 107-120 ‘2018.
6. Al-Rubaie, Fadel Mohsen, Abd, Salah Hamza, “Introduction to Chancery Operations”Directorate of Dar Al-Kutub for Printing and Publishing, (2005), Baghdad.
7. Al-Shamrti, Hamid and Al-Harthy, Abdul Rahim, “Random processes, their hypotheses and applications”, first edition, Baghdad, 2017.
8. Al-Shukurji ‘Thnoon Younis’ “Using hierarchical cluster analysis in classifying observations into homogeneous groups with an application to the basketball league” Journal of Research of the College of Basic Education ‘Volume 7’ No. 2 ‘pp. 335-347’ 2007.
9. Gavira-Durón, Nora, &, others, " Markov Chain K-Means Cluster Models and Their Use forCompanies’ Credit Quality and Default Probability Estimation", Mathematics 2021, 9, 879.
10. Habal’ Huda and Samaan ‘Jihad’, “Using multivariate analysis in studying the chemical, biological and industrial properties of wheat flour fortified with oatmeal.” Damascus Journal of Agricultural Sciences, Volume 28, Issue 2, pp. 361-374, 2012.
11. Hanish 'Dr. Ibrahim Salman and Asmayo 'Kholoud Suleiman' "Comparison between hierarchical clustering methods and choosing the best one with practical application on some types of milk sold in the city of Misurata" Special issue of the third annual conference on theories and applications of basic and biological sciences "Misrata University" Libya pp. 275-284 '2019.
12. Kazem ‘Feryal Muhammad’ and ‘Al-Alaq’ Dr. Mahdi Mohsen, “Statistical analysis of the Millennium Development Goals using the two methods of factor analysis and cluster analysis,” the Arab Institute for Training and Statistical Research, Baghdad, pp. 1-38, 2006.
13. Mustafa ‘Nizar’ “Using some methods of cluster analysis in classification with practical application” ‘Al-Taqni Journal’ Vol. 20’ No. 2 ‘Pg. 2’ 2007.
14. Nicolas Privault “Understanding Markov Chains Examples and Applications” Second Edition, Springer Nature Singapore Pte Ltd. 2018.
15. Taj, Lotfi and Ammar, Sarhan, "Introduction to Random Operations", First Edition, King Saud University, (2007), Riyadh, Saudi Arabia.