Who will be the FIFA World Champion in 2022 in Qatar

Zeiad Yahya Ali Allawee *College of Physical Education and Sports Sciences/University of Mosul/Iraq*
zead77@uomosul.edu.iq
Basel. Y. Thanoon *College of Computer Sciences and Mathematics University of Mosul. Iraq*
basytbb@yahoo.com

Received Date (31/08/2022) **Accepted Date (22/09/2022)**
DOI: (10.33899/rjss.2023.135629.1103)

ABSTRACT

In this paper, discrete-state Markov chains are used to obtain the winning probability through equilibrium and to estimate the Markov win probability value to predict the winners of the 2022 FIFA World Cup by modeling the historical participation of the participating teams in the 23rd FIFA World Cup which will be held in Qatar in November 2022.

We suggest a strength index based on some factors: the number of times participated in the World Cup, the number of times they reached the top four, and the number of times the team won first place, in addition to the factor of the ground and the audience (host.(

Predictive values are obtained for the teams that will top the tournament, which will qualify for the sixteenth round, which will qualify for the quarter-finals, semi-finals, and then the final Matches, which were arranged on the website of each team within its participation in the 2022 Napkin and according to the groups selected for it by FIFA.

Through this indicator, researchers were able to identify a model for participation in football. The teams within the rounds and their location in the round (16), the final, the quarter-finals, and to the final matches, gave us satisfactory and realistic results about the participation of each team and its role in the previous tournaments. It also gave us future space for an application to work on the rest of the competitions for team games such as handball, basketball, and volleyball.

Keywords : football, World Cup, Markov chains, probability.

من سيكون بطل العالم لكرة القدم عام ٢٠٢٢ في قطر؟

زياد يحيى علي علاوي كلية التربية البدنية وعلوم الرياضة / جامعة الموصل
zead77@uomosul.edu.iq
باسل يونس ذنون كلية علوم الحاسوب والرياضيات / جامعة الموصل
basytbb@yahoo.com

تاريخ تسليم البحث (٢٠٢٢/٨/٣١) **تاريخ قبول النشر (٢٠٢٢/٩/٢٢)**
DOI: (10.33899/rjss.2023.135629.1103)

الملخص

في هذا البحث، يتم استخدام سلاسل ماركوف ذات الحالة المنفصلة للحصول على احتمالية الفوز من خلال التوازن ولتقدير قيمة احتمالية فوز ماركوف للتنبؤ بالفائزين بكأس العالم لكرة القدم ٢٠٢٢ من خلال نمذجة المشاركة التاريخية للفرق المشاركة في بطولة العالم لكرة القدم الثالثة والعشرين. الكأس التي ستقام في قطر في نوفمبر ٢٠٢٢.

نقترح مؤشر قوة يعتمد على بعض العوامل: عدد مرات المشاركة في كأس العالم، عدد مرات الوصول إلى المراكز الأربعة الأولى، عدد مرات فوز الفريق بالمركز الأول، بالإضافة إلى عامل الأرض والجمهور (المضيف). .

يتم الحصول على القيم التنبؤية للفرق التي ستتصدر البطولة والتي ستتأهل إلى الدور السادس عشر والتي ستتأهل إلى الدور ربع النهائي ونصف النهائي ومن ثم إلى المباريات النهائية والتي تم ترتيبها على الموقع الإلكتروني الخاص بكل فريق ضمن مشاركتها في منديل ٢٠٢٢ وبحسب المجموعات المختارة لها من قبل الفيفا.

ومن خلال هذا المؤشر تمكن الباحثون من التعرف على نموذج للمشاركة في كرة القدم. الفرق ضمن الجولات ومواقعها في الدور (١٦) والنهائي وربع النهائي وصولاً إلى المباريات النهائية، مما أعطانا نتائج مرضية وواقعية حول مشاركة كل فريق ودوره في البطولة. البطولات السابقة. كما أتاح لنا مساحة مستقبلية للتطبيق للعمل على باقي المسابقات للألعاب الجماعية مثل كرة اليد وكرة السلة والكرة الطائرة.

الكلمات المفتاحية : كرة القدم، كأس العالم، سلاسل ماركوف، الاحتمالية.

1 - Introduction

Football is one of the most famous and famous university sports in the world, It has become important, especially in the World Cup tournament from its first edition in the year 1930 until now, and it has been the focus of all countries' attention to achieving victory. They were participating in this global tournament because of its importance at all economic and political levels and for the development of countries that worked with all their energies to gain residency in their country, as the number of participants in its last two editions reached 32 teams, which ended after the qualifying stage, which included more than 200 teams.

The Brazilian team won five championships, the only team that participated in all its 21 editions, and the championship was won by teams most of them from the European continent (FIFA).

Eight countries won the World Cup, as we all mentioned from the two continents (South America and Europe) in the year 1966. The English team had its first and only victory in the championship, and then other European teams such as (France and Spain) followed the list of winners of the championship, and no team from Africa or Asia was honored to be crowned despite the strength of African teams such as (Nigeria, Ghana and Cameroon), and on the other hand, there are Asian teams that participated more than once to be nominated for the World Cup finals and did not get the championship, such as (Australia, Turkey, Japan and South Korea) despite their strength and skill.

Many researchers and analysts have been interested in conducting a lot of research related to the World Cup and predicting the results of the teams and who will get the championship in different continents.

In the year (2007) Lago presented a study on the impact of performance on the results of teams in the 2006 World Cup, in which he found the importance of performance during the first rounds of competitions and the effect of performance in the knockout stages became less important, and whose analyzes were based on multiple regression and comparison test for averages and the polynomial logarithm model.

Paul and Mitraj (2008) used two empirical models to analyze four winners of the World Cup through the monthly FIFA ranking, which confirmed that all teams ranked within the first place were the winners of the World Cup except in 1994 by examining the strength of the uncertainty element.

In the year (2009) VOLF presented a near-parametric double regression model to predict the results of the qualification for the 2006 World Cup quarter-finals.

Kaufmann presented in the year (2014) a proposal that confirmed that good governance in the country and the fan base with awareness, high culture and love of support for the team are among the most important factors that help teams win the World Cup using the multiple linear regression model for the variables of the size of the economy of the country and the market value of the coach the team (Kaufmann, 2014).

Seth (2018) introduced a multiple linear regression model in which he examined the determinants of FIFA World Cup team performance. One of the important variables in this model is the team's ranking within the tournament's classification, the presence of professional players and stars in the team, in addition to whether the country is hosting the World Cup.

Ikoba (2021) Markovian modeling was used to predict the winner of the FIFA World Cup in the World Cup 2022, which showed the reliability of the model by forming the initial transition matrix and calculating the probability of winning through the results obtained by the participating teams in all previous World Cup matches.

Researchers Ziad and Thanoun (2022) also presented a proposed Markovian model to predict the teams most likely to win the World Cup title held in Qatar in 2022 by modeling the historical matches of all participating teams, forming a Markovian matrix, and obtaining a vector of wins, draws, and losses, which gave close and satisfactory results from by matching the results that occurred in the tournament. In addition to creating a strength index, researchers Ziad and Thannoun (2023) can draw a qualification map using an indicator called the strength index to win the World Cup.

In Section 2, the methodology was described to describe the Markov chains model, its theoretical analysis, and the proposed model for team strength. In the third section, the results of the analysis of Markov chains were reviewed, the formation of the transition matrix for each team, the determination of the value of the equilibrium vector, the presentation of the data of the teams participating in the 2022 World Cup,

and the identification of the teams that would You get to the round of 16, the semi-finals, the quarter-finals, the final match and the selection of the top four, and finally, the conclusions are expressed in Section 4.

2- The Problem:

Football has become a difficult digital factor, despite the simplicity and ease of the game, but many fans of the game and specialists in the sports field, as well as officials in charge of supervising and coaching teams and pawnbrokers, are looking for multiple ways to predict the team that wins the championship or competition, especially international championships such as the World Cup and the European Championships for national teams and others at all levels, even salty ones.

The development of the football game and the development of the teams' performance largely and largely, despite the increase in the number of teams participating in this tournament (the World Cup Championship) to 32 teams playing for the finals after more than 200 teams were eliminated through the qualifiers stage and access to the finals. On the one hand, the big role that the media plays in granting this primary tournament and the countries that win hosting it and winning the title. All of these factors push us to find out the use of statistical methods and models and mathematical models that help us create a distinct predictive model that helps us analyze the data of football sports teams and work to find a champion crowned in this tournament by finding A numerical index for the purpose of giving first place to the winning team.

2-1 The purpose of the research

It is finding a numerical index (evidence) as a mathematical model using a random Markov chain to predict the winner of the 2022 World Cup that will be held in Qatar. A comprehensive analysis is presented to predict the future match based on steady-state probabilities.

2-2 Research Areas

- Human field: the official national teams nominated for the 2022 World Cup final in Qatar.

- Historical field: The results of the official matches of the teams participating in the 2022 World Cup in terms of their participation and the results obtained by these teams in all World Cup tournaments from 1930 to 2018.

2-3 Search terms:

- Variables: The characteristics and advantages that show a difference in their values are called variables, and there are two types of variables:

- Descriptive (Qualitative) variables:

They are the variables that represent specific characteristics and characteristics that are measured descriptively (not including numerical significance) such as color, gender, and DNA.

- Quantitative Variables

They are the variables that have a numerical value resulting from a numerical measurement of something, as a specific physical aspect. These variables are either continuous (continuous) and usually take values in the form of real numbers, such as: height, weight ... etc. or discrete (discrete) and usually It takes values in the form of integers (positive or negative), such as number of balls, number of goals, number of pitches etc.

- State space:

It is defined as the set of all possible values of the stochastic process called the state space and symbolized by S.

The state space can be finite or infinite depending on the number of its elements. Also, it can be continuous if it takes continuous values, or discrete if it takes separate values.

- Markov processes:

The stochastic process {X(t);t∈T} is said to be a Markovian process if the values are

empty ass $t_1 < t_2 < \cdots < t_n < t$ as follows:

$P\{a \le X_n \le b/X(t_1) = X_n. \ldots . X(t_n) = X_n\}=$
$P\{a \le X_n \le b = X(t_n) = X_n\}$ (1)

It is called a Markov chain if the state space is discrete, which is often in the form of a set of integers {0.1.2.…} as the state space of the Markov chain Thanoon (2011).

.It is called a continuous Markov process if the state space is continuous which is in the form

$$-\infty < t < \infty$$

- Markov chains:

Markov chains are a special type of Markov process that can be represented by the discrete state space and the discontinuous and continuous parameter space, and they are a sequence of random variables that fulfill the Markov property, and the systems that have this property are called Markov chains.

3- Research Methodology:

In this section, we focus our attention on stochastic processes with discrete state space in discrete time. For convenience, we will symbolize the state space elements, regardless of their names, with the integer {1,2,3,} (with some special exceptions).

If the discrete random variable represents the "state" of an observed system at the time n .

As is common in specialized scientific sources, we will symbolize the stochastic process with a discrete state space with the symbol $\{X_n; n = 0,1,2,...\}$ or for short $\{X_n\}$

Usually, the n index stands for time, because this index is the most commonly used in applications, however can be any other index, such as the location somewhere. Since the index n represents time, "n" represents the present, "n-k" represents the past before k units of time, and "n+k" represents the future after k units of time. The observed "data" is actually from the stochastic process, so let it be that the stochastic process $\{X_n\}$ is said to have the Markovian property if it $X_0, X_1,..., X_n,...$ achieves the following:

$$P(X_n = j / X_{n-1} = i, X_{n-2} = a, X_{n-3} = b,..., X_0 = c)$$
$$= P(X_n = j / X_{n-1} = i) = p_{ij,}$$

For all $.n, i, j, a, b, c$. Gagniuc (2017).

The stochastic process has a discrete state space and bears the Markovian property.

The transition possibilities from state i to state j are called transitional probabilities and are symbolized by them p_{ij}. If the transition probabilities do not depend on time n then the Markov chain is described as time-homogeneous, or stationary.

When the state space S is finite, the probability distribution of transitions between states can be represented by a matrix called the Transition Matrix, and is symbolized by it P ,as the element (i, j) of this matrix represents the transition probability $p_{ij} = P(X_n = j / X_{n-1} = i)$. (2)

A transition matrix P is a stochastic matrix that satisfies the following two conditions:

1- All of its elements are non-negative (because they are probabilities).

2- The sum of each of its rows is 1 (because the sum of the total probabilities is equal to 1).

If the state space is in the form of S= {1,2,…, N}, then the general form of a one-step transition matrix is as follows:

$$P = \begin{array}{c} \\ X_{n-1}=1 \\ 2 \\ . \\ N \end{array} \begin{array}{c} \begin{array}{cccc} X_n=1 & 2 & ... & N \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ . & . & \cdots & . \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix} \end{array}. \quad (3)$$

The matrix $P = (Pij)$ is the string transition matrix Thanoon (2011).

As we mentioned, the transition possibilities from state i to state j, p_{ij} are called transitional probabilities, meaning that the transition probability is the conditional probability $p_{ij} = P(X_1 = j / X_0 = i)$. Or more precisely, p_{ij} the probability of moving from state i to state j is one-step. The probability of moving from state i to state j in n steps is denoted by it $p_{ij}^{(n)}$, and it is clear that:

$$p_{ij}^{(n)} = P(X_n = j / X_0 = i); n = 1,2,...$$

It is noted that $p_{ij}^{(1)} = p_{ij} = P(X_1 = j / X_0 = i)$, and

$$p_{ij}^{(2)} = P(X_2 = j / X_0 = i) = \sum_{r \in S} P(X_2 = j, X_1 = r / X_0 = i)$$

$$= \sum_{r \in S} P(X_2 = j / X_1 = r, X_0 = i) P(X_1 = r / X_0 = i)$$

$$= \sum_{r \in S} P(X_2 = j / X_1 = r) P(X_1 = r / X_0 = i)$$

According to the Markovian property

$$\therefore p_{ij}^{(2)} = \sum_{r \in S} P(X_1 = r / X_0 = i) P(X_2 = j / X_1 = r)$$

$$= \sum_{r \in S} p_{ir} p_{rj.}$$

As for the transition matrix with k steps, $P^{(k)}$, this matrix represents the probability of moving from state I to state j in k steps, that is, the elements of this matrix are $p_{ij}^{(k)} = P(X_n = j / X_{n-k} = i)$. (4)

Axiom: Thanoon (2011)

For any $o < k < n$, the transitional probability of n steps from state i to state j satisfies the following equation:

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}.$$

the proof:

$$p_{ij}^{(n)} = P(X_n = j / X_0 = i)$$

$$= \sum_{r \in S} P(X_n = j, X_k = r / X_0 = i)$$

Because the string must be in some state, say r, at time k.

$$\therefore p_{ij}^{(n)} = \sum_{r \in S} P(X_n = j / X_k = r) P(X_k = r / X_0 = i)$$

Sines $P(A \cap B/C) = P(A/B \cap C)P(B/C)$.

That is:

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}.$$

The series must have a steady state (or steady-state) distribution if there is a vector like this given the transition probability matrix P:

$$\pi = \pi p \qquad (5)$$

If the finite Markov chain is irreducible, then:$\lim_{n\to\infty} p^n = \pi = \begin{bmatrix} \pi_1 & \cdots & \pi_n \\ \vdots & \ddots & \vdots \\ \pi_1 & \cdots & \pi_n \end{bmatrix}$ (6)

where $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ with $0 < \pi_j < 1$ and $\Sigma\pi_j = 1$

This constant probability vector can be viewed as the only long-run distribution of a random variable.

Furthermore, the steady-state probabilities of π_j are obtained as$\lim_{n\to\infty} {p_{ij}}^n \; \forall i.j \in S$ Thanoon (2011).

4- Building a Markov Chain Model

If we give a series of observed states (verified observations), the first thing to note is that these observations may be in the form of symbols of two or more letters, or integers, usually in the form of 1.2 ... n. To facilitate computer manipulation, it is preferable to convert literal symbols into integers. If we assume that the observations achieved are: {a d b a a c b d a d a c d b a d c d b d a d c c a b}

These symbols can be encoded into integers as follows: a = 1, b = 2, c = 3, d = 4. It is possible to take advantage of the following program that encodes any combination of letters or symbols into integers, and this program is based on observations The previously verified results, which represent the results obtained by the teams through their participation in the tournaments held by FIFA called the World Cup from 1930 to 2018, in addition to the results obtained by these teams from 2018 until the last match in 2022 before the establishment of the World Cup, which was obtained On them through the website (Wikipedia, the free encyclopedia), the number of transitions from each state to another was obtained and the transition probabilities were estimated using the maximum likelihood criterion.

The following table shows the total number of matches played by each team during its participation in the World Cup finals, the number of victories, the number of

draws, and the number of losses for each team included in this study, which was arranged on the website of each team within its participation in the 2022 Napkin and according to the groups selected for it by FIFA with the number of hosting times For the championship, winning the cup, the number of times to qualify for the World Cup finals, and the number of times to reach the four advanced positions.

Table (1): number of matches played by each team during its participation in the World Cup finals, the number of victories, the number of draws, and the number of losses for each team included in this study, which were arranged on the website of each team within its participation in the 2022 Napkin and according to the groups selected for it by FIFA with the number of hosting times For the championship, winning the cup, the number of times to qualify for the World Cup finals, and the number of times to reach the four advanced positions.

Table 1

Group	Team	number of matches	win	Draw	loss	The four advanced positions	Number of times hosting the tournament	Number of times to participate in the World Cup	Number of times to win the World Cup
First	Qatar	0	0	0	0	0	1	0	0
	Ecuador	10	4	1	5	0	0	3	0
	Senegal	5	2	2	1	1	0	2	0
	Netherlands	50	27	12	11	5	0	10	0
Second	England	62	26	20	16	3	1	15	1
	Iran	12	1	3	8	0	0	5	0
	United States	33	8	6	19	1	1	10	0
	Wales	5	1	3	1	0	0	1	0
Third	Argentina	77	42	14	21	5	1	17	2
	Saudi Arabia	13	2	2	9	0	0	5	0
	Mexico	53	14	14	25	0	2	16	0
	Poland	31	15	5	11	2	0	8	0
Fourth	France	59	28	12	19	6	2	15	2
	Austria	13	2	3	8	0	0	5	0
	Denmark	16	8	2	6	0	0	5	0
	Tunisia	12	1	4	7	0	0	5	0
Fifth	Spain	59	29	12	18	2	1	15	1
	Costa Rica	15	5	4	6	0	0	5	0
	Germany	106	66	20	20	12	2	19	4
	Japan	17	4	4	9	0	1	6	0
Sixth	Belgium	41	14	9	18	0	0	13	0
	Canada	3	0	0	3	0	0	1	0
	Morocco	13	2	4	7	0	0	5	0
	Croatia	16	7	2	7	1	0	5	0
Seven	Brazil	104	70	17	17	11	2	21	5
	Serbia	43	17	8	18	2	0	12	0
	Switzerland	33	11	6	16	11	1	11	0
	Cameroon	23	4	7	12	0	0	7	0
Eight	Portugal	26	13	4	9	1	0	7	0
	Ghana	12	4	3	5	0	0	3	0
	Uruguay	51	20	12	19	5	1	13	2
	South Korea	31	5	9	17	1	1	10	0

To build a Markov chain model, Xn is the team's performance on the last day and is named as

State 1: The team wins the match (W).

State 2: The team is drawn in the match (D).

State 3: The team loses the match (L).

So, $\{ Xn \}$ is a random process that has a value from 1 to 3 in states. Furthermore, consider the team's performance in any sequence of letters chosen from the combination (W,D,L). Suppose $n_1 . n_2 . n_3$ denotes the number of letters "W", "D" and "L"; respectively in the sequence.

Frequencies Matrices for transitions from one state to another:

The three cases are here and this sequence of results is recorded in a

3×3 matrix.

$$F = \begin{array}{c} \\ W \\ D \\ L \end{array} \begin{array}{c} \begin{array}{ccc} W & D & L \end{array} \\ \begin{bmatrix} n_{ij} & n_{ij} & n_{ij} \\ n_{ij} & n_{ij} & n_{ij} \\ n_{ij} & n_{ij} & n_{ij} \end{bmatrix} \end{array} \qquad (7)$$

The transition probability matrix P was constructed by dividing each element by the corresponding row total.

$$\text{W} \quad P = P_{ij} = \begin{array}{c} W \\ D \\ L \end{array} \begin{bmatrix} P_{ij} & P_{ij} & P_{ij} \\ P_{ij} & P_{ij} & P_{ij} \\ P_{ij} & P_{ij} & P_{ij} \end{bmatrix} \qquad (8)$$

With $\sum_j P_{ij} = 1$, for all i.

This transition probability matrix is called the $[3 \times 3]$ m rix of the transition probability matrix of the first step of the Markov chain. Moreover, each row of P is the probability distribution related to the transition from state i to state j.

Further, te steady-state distribution is examined and ies are obtained basebtaied$\pi_j = \lim_{n\to\infty} {p_{ij}}^n \; \forall i.j \in S$

In light of this interpretation of the stationary distribution, we propose the following mathematical indicator of historical mathematical excellence for the probability of winning; $SCUP_Q$; As a final summary of the historical results of the sports teams participating in the World Cup finals from 1930 until 2018, with the results they obtained after 2018, considering a stage close to the qualifiers, in addition to the fact that there are teams that did not participate in the World Cup finals and there are teams that only participated once, which Through this, it is not possible to form an iterative matrix for the results, as well as a transitional matrix for them.

The World Cup winning factor (F) has been added as it has a significant impact on the matches Lago (2007), as well as the factor of reaching the top four places (R) Suzuki et al. (2010) and the hosting factor (audience and ground) (E) Cattelan, et al. (2013) . As well as the filter factor for the World Cup finals (S) Paul and Mitra (2008), all of these factors were given a weight restiform multiplying the number of times by a numerical value equal to 0.1, so he index of excellence team will take becomes as follows:$SCUP_Q = \pi_w + \mu(F) + \mu(R) + \mu(E) + \mu(S)$ (9)

Where:$\pi_w = The\ balanced\ value\ ofofe\ probability\ of\ winning$

$\mu =$ The number of times.

5- Applying Markovian Modeling:

We consider the result of any sporting match as states of a 3- stats Markov chain with states: Win (W), a Draw (D), or a Loss (L).

The basic assumption in the Markovian model is that the result of the current match depends on the result of the previous match, and that the result of that that preceded it depends on the outcome of its previous match, and the result of that previous match depends on the outcome of the match that preceded it and so on; see e.g.

If {X1, X2,…} represents the results of consecutive matches for a specific team, then the probabilistic model expressing the Markovian property can be formulated mathematically as follows; see e.g. Gagniuc (2017):

$$P(X_n|X_{n-1}.X_{n-2}.\ldots) = P(X_n\,|X_{n-1}); n = 1.2.\ldots \quad (10)$$

Table (2) shows the frequencies observed when moving from one state to another in successive matches; for example, WD stands for the transition from a win state to a draw state in two consecutive matches, while DW stands for the transition from a draw state to a win state in two consecutive matches.

To clarify understandably, Table No. (2) was prepared according to the frequency matrix, which was referred to in the equation (7).

Table (2)

The observed frequencies of states for the selected teams

Group	Team	WW	WD	WL	DW	DD	DL	LW	LD	LL
first	Qatar	20	4	5	4	0	3	4	4	6
	Ecuador	8	4	5	4	5	5	6	6	7
	Senegal	22	8	7	8	3	2	7	2	0
	Netherlands	30	9	12	10	2	3	12	4	2
second	England	33	9	18	13	6	4	13	8	6
	Iran	17	2	6	3	1	1	5	2	7
	United States	24	8	9	7	2	4	9	4	7
	Wales	10	4	7	7	5	2	4	4	2
third	Argentina	40	14	19	19	2	1	14	6	7
	Saudi Arabia	13	3	8	4	1	5	7	6	10
	Mexico	16	13	17	10	3	9	20	7	18
	Poland	20	6	12	7	2	4	11	5	6
fourth	France	41	9	13	9	3	6	12	6	10
	Austria	12	2	4	3	1	4	3	5	8
	Denmark	18	7	9	6	3	3	10	2	3
	Tunisia	13	5	10	4	2	6	11	5	4
Fifth	Spain	32	12	10	13	5	5	9	6	6
	Costa Rica	8	4	5	4	4	5	5	5	13
	Germany	53	12	21	16	8	3	20	4	3
	Japan	22	4	10	6	1	3	9	5	6
Sixth	Belgium	32	7	8	5	1	3	11	1	8
	Canada	19	1	8	2	2	0	7	1	2
	Morocco	21	6	6	8	4	3	5	5	7
	Croatia	15	5	11	8	1	4	8	7	7
Seven	Brazil	58	11	14	11	1	2	15	2	1
	Serbia	9	7	2	5	1	3	4	2	8
	Switzerland	11	6	12	9	2	7	9	10	11
	Cameroon	12	6	7	9	4	4	5	6	9
eight	Portugal	24	8	7	8	2	4	6	4	3
	Ghana	4	3	8	5	3	4	6	6	4
	Uruguay	27	9	12	11	3	7	10	9	11
	South Korea	19	8	6	7	3	6	8	5	12

The elements of the frequency matrices of the teams have been transformed into transition matrices of probabilities; each with a dimension 3 × 3.

After obtaining the matrix of frequencies, we calculate the transition matrix, this program is called MCestimation Thanoon (2011), as this program finds both the matrix of frequencies as well as the transition matrix for any observations made from a Markovian chain consisting of S from the possible cases, and according to the definition of the relative frequency of the probability, that We form the transition matrix by dividing each element in the frequency matrix by the sum of the row in which it is located. That is, the transitional probabilities are calculated as follows:

Table (3)
Elements of transition probability matrices for the selected teams

group	the team	WW	WD	WL	DW	DD	DL	LW	LD	LL
first	Qatar	0.6897	0.1379	0.1724	0.5714	0	0.4286	0.2857	0.2857	0.4286
	Ecuador	0.4706	0.2353	0.2941	0.2667	0.3333	0.4	0.3158	0.3158	0.3684
	Senegal	0.5946	0.2162	0.1892	0.6154	0.2308	0.1538	0.7778	0.2222	0
	Netherlands	0.5882	0.1765	0.2353	0.6667	0.1333	0.2	0.6667	0.2222	0.1111
second	England	0.55	0.15	0.3	0.5652	0.2609	0.1739	0.4815	0.2963	0.2222
	Iran	0.68	0.08	0.24	0.6	0.2	0.2	0.3571	0.1429	0.5
	United States	0.5854	0.1951	0.2195	0.5385	0.1538	0.3077	0.45	0.2	0.35
	Wales	0.4762	0.1905	0.3333	0.5	0.3571	0.1429	0.4	0.4	0.2
third	Argentina	0.5479	0.1918	0.2603	0.8636	0.0909	0.0455	0.5185	0.2222	0.2593
	Saudi Arabia	0.5417	0.125	0.3333	0.4	0.1	0.5	0.3043	0.2609	0.4348
	Mexico	0.3478	0.2826	0.3696	0.4545	0.1364	0.4091	0.4444	0.1556	0.4
	Poland	0.5263	0.1579	0.3158	0.5385	0.1538	0.3077	0.5	0.2273	0.2727
fourth	France	0.6508	0.1429	0.2063	0.5	0.1667	0.3333	0.4286	0.2143	0.3571
	Austria	0.2222	0.4444	0.3333	0.375	0.125	0.5	0.428571	0.214286	0.357143
	Denmark	0.5294	0.2059	0.2647	0.5	0.25	0.25	0.6667	0.1333	0.2
	Tunisia	0.4643	0.1786	0.3571	0.3333	0.1667	0.5	0.55	0.25	0.2
Fifth	Spain	0.5926	0.2222	0.1852	0.5652	0.2174	0.2174	0.4286	0.2857	0.2857
	Costa Rica	0.4706	0.2353	0.2941	0.3077	0.3077	0.3846	0.2174	0.2174	0.5652
	Germany	0.5955	0.1685	0.236	0.5926	0.2963	0.1111	0.7407	0.1481	0.1111
	Japan	0.6111	0.1111	0.2778	0.6	0.1	0.3	0.45	0.25	0.3
Sixth	Belgium	0.6809	0.1489	0.1702	0.5556	0.1111	0.3333	0.55	0.05	0.4
	Canada	0.6786	0.0357	0.2857	0.5	0.5	0	0.7	0.1	0.2
	Morocco	0.6364	0.1818	0.1818	0.5333	0.2667	0.2	0.2941	0.2941	0.4118
	Croatia	0.4839	0.1613	0.3548	0.6154	0.0769	0.3077	0.3636	0.3182	0.3182
Seven	Brazil	0.6988	0.1325	0.1687	0.7857	0.0714	0.1429	0.8333	0.1111	0.0556
	Serbia	0.5	0.3889	0.1111	0.5556	0.1111	0.3333	0.2857	0.1429	0.5714
	Switzerland	0.3793	0.2069	0.4138	0.5	0.1111	0.3889	0.3	0.3333	0.3667
	Cameroon	0.48	0.24	0.28	0.5294	0.2353	0.2353	0.25	0.3	0.45
eight	Portugal	0.6154	0.2051	0.1795	0.5714	0.1429	0.3857	0.4615	0.3077	0.2308
	Ghana	0.2667	0.2	0.5333	0.4167	0.25	0.3333	0.375	0.375	0.25
	Uruguay	0.5625	0.1875	0.25	0.5238	0.1429	0.3333	0.3333	0.3	0.3667
	South Korea	0.5758	0.2424	0.1818	0.4375	0.1875	0.375	0.32	0.2	0.48

In Table (4), the possibility of the transitional matrix was clarified for each team to facilitate the discussion process and according to the conditions of the transitional matrix in equation (8).

6- A Suggested Index:

In this indicator, we took the balanced value of the probability of winning for each team plus the factor of winning the World Cup multiplied by the number of times the team won the championship, plus the number of times the team qualified for the World Cup final, the number of times hosting the World Cup and the number of times reaching the four advanced positions , some properties of this proposed mathematical indicator can easily be deduced:

1- Theoretically, its highest value is 100% when all the results of the team are wins without any draw or loss.

2- The indicator signal is a clear indication of the positive or negative results of the team.

3- When comparing historically between a group of teams, the team with the highest value of this indicator is the best and it will be the champion of Mendel 2022.

4 - The teams are compared by giving each team the index that it obtained, choosing the best from each group, determining the second winner for each group, and then nominating them for the 16th round, during which the teams will compete according to their strength until reaching the final matches and determining the World Cup champion.

Note that when calculating this index for a group of teams, it is best to normalize it by dividing the value of each team's index by the value of their largest index and then multiplying the result by 100%.

6.2 Results and Discussion:

Using the transition probability matrices for the selected teams shown in Table (3), the fixed probability distribution of the gains was obtained for each team with the weights of the factors mentioned in Table No. (1) as shown in Table (4).

Table (4)

Fixed probability distribution of winnings, weight of the World Cup winning factor W_F, the number of times to qualify for the World Cup finals W_R , the number of times to host W_E , and access to the four advanced positions for each team W_S.

Group	Team	π_w	W_F	W_S	W_E	W_R
First	Qatar	0.5553	0	0	0.1	0
	Ecuador	0.3567	0	0.3	0	0
	Senegal	0.6272	0	0.2	0	0.1
	Netherlands	0.6182	0	1	0	0.5
Second	England	0.5358	0.1	1.5	0.1	0.3
	Iran	0.5682	0	0.5	0	0
	United States	0.565	0	1	0.1	0.1
	Wales	0.4644	0	0.1	0	0
Third	Argentina	0.5984	0.2	1.7	0.1	0.5
	Saudi Arabia	0.4211	0	0.5	0	0
	Mexico	0.4071	0	1.6	0.2	0
	Poland	0.5205	0	0.8	0	0.2
Fourth	France	0.5662	0.2	1.5	0.2	0.6
	Austria	0.4286	0	0.5	0	0
	Denmark	0.5574	0	0.5	0	0
	Tunisia	0.4667	0	0.5	0	0
Fifth	Spain	0.551	0.1	1.5	0.1	0.2
	Costa Rica	0.3208	0	0.5	0	0
	Germany	0.6218	0.4	1.9	0.2	1.2
	Japan	0.5631	0	0.6	0.1	0
Sixth	Belgium	0.6336	0	1.3	0	0
	Canada	0.6667	0	0.1	0	0
	Morocco	0.5302	0	0.5	0	0
	Croatia	0.4697	0	0.5	0	0.1
Seven	Brazil	0.7294	0.5	2.1	0.2	1.1
	Serbia	0.4479	0	1.2	0	0.2
	Switzerland	0.3766	0	1.1	0.1	1.1
	Cameroon	0.4184	0	0.7	0	0
Eight	Portugal	0.5732	0	0.7	0	0.1
	Ghana	0.3489	0	0.3	0	0
	Uruguay	0.4848	0.2	1.3	0.1	0.5
	South Korea	0.4642	0	1	0.1	0.1

From the last table, the sports index values for the historical sports excellence index; $SCUP_Q$; were calculated and presented in the following table with the corresponding measured values.

alterable (5)$SCUP_Q$values for the selected football teams and corresponding normal values.

Group	Team	$SCUP_Q$	Normalized
First	Qatar	0.6553	14.15518
	Ecuador	0.6567	14.18542
	Senegal	0.9272	20.02851
	Netherlands	2.1182	45.75539
Second	England	2.5358	54.776
	Iran	1.0682	23.07426
	United States	1.765	38.12589
	Wales	0.5644	12.19164
Third	Argentina	3.0984	66.92876
	Saudi Arabia	0.9211	19.89675
	Mexico	2.2071	47.67572
	Poland	1.5205	32.84443
Fourth	France	3.0662	66.23321
	Austria	0.9286	20.05875
	Denmark	1.0574	22.84097
	Tunisia	0.9667	20.88176
Fifth	Spain	2.451	52.94423
	Costa Rica	0.8208	17.73016
	Germany	4.3218	93.35551
	Japan	1.2631	27.28431
Sixth	Belgium	1.9336	41.76783
	Canada	0.7667	16.56154
	Morocco	1.0302	22.25342
	Croatia	1.0697	23.10667
Seventh	Brazil	4.6294	100
	Serbia	1.8479	39.91662
	Switzerland	2.6766	57.81743
	Cameroon	1.1184	24.15864
Eighth	Portugal	1.3732	29.66259
	Ghana	0.6489	14.01694
	Uruguay	2.5848	55.83445
	South Korea	1.6642	35.9485

It is clear from the results obtained from this proposed historical indicator of the strength of the teams that will participate in the 2022 World Cup, which is held in the State of Qatar, that the Brazilian team will get first place and win the championship for the sixth time, and how much we can give a clearer picture of the teams that It will be nominated for the round (16), which are: (Netherlands and the United States), (Argentina and Denmark), (Germany and Croatia), (Brazil and South Korea), (England and Senegal), (France and Mexico), (Belgium and Spain), (and the Uruguayan team). And Switzerland) and the teams that will compete in the quarter-finals, which are: (Netherlands and Argentina), (Germany and Brazil), (England and France), (the Spanish team and Switzerland), and the teams that will advance to the semi-finals are ((Argentina and Brazil), (and France and Switzerland)) and the teams that will occupy the four advanced positions, which are (Brazil in first place, France

in second place, and Argentina in third place, and fourth place will be from the Swiss team's share).

To clarify the strength of the historical index of the teams participating in the 2022 World Cup, we can infer the classification of the teams after the end of the tournament and as it is in force in FIFA Figure (1): Classification of the selected teams according to the proposed sports index for historical sports excellence.

7- Conclusions and recommendations

1- We presented a predictive model of who will win the 2022 World Cup, which will be held in Qatar in October, by modeling the historical matches of the teams that qualified for the 2022 World Cup through their historical participation in previous tournaments held by FIFA from 1930 to 2018 using discontinuous Markov chains, which are from During which we obtained the balanced value of the winning cases.

2 - We analyzed the results of 32 teams according to their potential as a ranking by giving an indicator that reflects the current strength of the team. Our analysis shows that Markov chain models gave the advantage of not ignoring the historical activity of the teams, which is the basis of the team's superiority through the match result in terms of winning.

3 - We gave importance to the teams that participated in the World Cup for the previous tournaments, in addition to highlighting the factor of the public and the land and their importance in determining the outcome of the matches and gave equal weight to the teams that had many participations in the World Cup with the strength of the teams that reached the four advanced positions.

4 - We found in this paper an easy and very important way to find out who will be the World Cup champion in 2022, and how it can be exploited by scheduling the matches played by each team. We showed through our experiences how to seize opportunities to move to advanced positions in the championships. And opportunities to advance in the table in terms of accompanying possibilities.

3- Our proposed methods can be used to predict which teams will win the championships held at the continental, national, and local levels.

4- We also benefited from betting odds as a highly valuable tool in processing the available information and predicting sporting events. The odds of the bet themselves are a measure of the expectations of success in the next match and using our approach we can relate these market expectations directly to the quantitative rating of each team, i.e. a measure of team quality.

6- We can also use our suggestions for all team and individual games such as basketball, volleyball, handball...etc, among the games followed by people, analysts, etc.

- References:

[1] FEDERATION INTERNATIONALE DE FOOTBALL FEDERATION (FIFA). URL:www.fifa.com

[2] Gagniuc, Paul A. (2017). Markov Chains: From Theory to Implementation and Experimentation. USA, NJ: John Wiley & Sons. pp. 1–256. ISBN 978-1-119-38755-8.

[3] Ikoba, N. A. A (2021): Sequential Markov Chain Model of FIFA World Cup Winners.

[4] LAGO, C. (2007). Are Winners Different from Losers? Performance and Chance in the FIFA World Cup Germany 2006. International Journal of Performance Analysis in Sport, 7 (2), 36 – 47.

[5] PAUL, S. AND MITRA, R. (2008). How Predictable are the FIFA World Cup Football Outcomes? An Empirical Analysis. Applied Economics Letters, 15, 1171 – 1176.

[6] SETH, S. (2018). FIFA World Cup: Factors that explain the performances of National Football Teams. Claremont Colleges Thesis. http://scholarship.claremont.edu/cmc_theses/1919.

[7] Thanoon, B.Y. (2011). "Markovian Modeling with Practical Applications", Part One, Ibn Al-Atheer House for Printing and Publishing, University of Mosul.

[8] VOLF, P. (2009). A random point process model for the score in sports matches. IMA Journal of Management Mathematics 20(2): 121–131.

[9] WIKIPEDIA. www.wikipedia.org/FIFA_World_Cup.

[10] SUZUKI, A. K.; SALASAR, L. E. B.; LEIT, J. G. AND LOUZADA-NETO, F. (2010). A Bayesian Approach for Predicting Match Outcomes: The 2006 (Association) Football World Cup. The Journal of the Operational Research Society, 61 (10), 1530 – 1539.

[11] CATTELAN, M.; VARIN, C. AND FIRTH, D. (2013). Dynamic Bradley-Terry Modeling of Sports Tournaments. Journal of The Royal Statistical Society, Series C (Applied Statistics), 62 (1), 135 – 150.

[12] Ali Allawee, Z. Y., & Thanoon, B. Y. (2022). Markovian modeling of match results with an application. Journal of Interdisciplinary Mathematics, 25(8), 2671-2677.

[13] Allawee, Z. Y. A., & Thanoon, B. Y. (2023). An Index of Superiority for the Most Famous International Football Teams. *Al-Rafidain Journal For Sport Sciences*, *26*(80), 448-461.