Problem Chosen	2023 MCM/ICM Summary Sheet	Team Control Number
Z		2330619

Who are Taking Up the Bidding Baton?

The objective of this paper is to identify the key factors that impact a country's willingness to bid for hosting the Olympic Games and quantify their effects. Drawing from our analysis, we offer actionable recommendations to foster greater interest in bidding for the Olympics and ensure the successful hosting of the games.

Firstly, We **calculated the application information** of countries that have applied to host the Olympic Games and used the **average shift method** to obtain the bid willingness of each bidding country in different years.

Secondly, We collected two types of data to support our study. **The first part** is the data related to the host cities that have hosted the Summer Olympics Games or Winter Olympics Games. Five indicators were selected: Contribution to GDP and tourism, Venue abandonment rate, public satisfaction and impact, and ROI. Then we utilize the TOPSIS weighed by **entropy method** to score the previous Olympic Games. And **the second part** is comprehensive data on the **world's development** over the past seven decades, including changes in economic, social, and investment in education and culture and international influence. From these, we extracted the changes in the level of national development of the countries bidding to host the Olympic Games.

Thirdly, In order to better analyze the influence of factors such as **the level of development** of a country on its willingness to bid for the Olympics, we used **cluster analysis** to classify all bidding countries into four categories using ward distances according to the data used in the previous section.

Fourthly, We exert **factor analysis** to identify the most siginigicant factor influencing the willingness of each category of countries. We considered four public factors: **empirical** factors(the effect of hosting the previous two Olympic Games), **economic** factor, **social** factor, cultural and **educational** factor, and one special factor: international political and cultural influence factor. After analysis, we believe that the willingness of countries in the fourth category to host the Summer Olympics in the next two sessions is the strongest, followed by the second category; the willingness of countries in the third category to host the Winter Olympics is the strongest, followed by the fourth category. This reveals that **developed countries are still the major players in bidding for the Olympic Games**, but with the increase in bidding costs and world economic and social development, the baton of bidding for the Olympic Games is being **handed over to countries that are rapidly developing** and actively seeking international influence

Fifthly, We performed validation and sensitivity analysis of the previously established model to calculate the willingness level based on the data of 2024 as well as 2026 Olympic Games bidders and other non-bidding countries. The calculation results are consistent with the actual situation. Then we took the data from previous years to conduct sensitivity analysis, and the analysis results verified the correctness of the model.

Finally, we synthesize each indicator and its importance and give suggestions on how to improve the motivation of countries to bid for the Olympic Game. We conducted accuracy testing and sensitivity analysis on the model, and it works.

Keywords: Olympic Games TOPSIS Cluster Analysis Factor Analysis willingness

Content

1 Introduction

1.1 Problem Background

In the past, the competition for hosting the Olympic Games has been fiercely contested. However, due to various short-term and long-term negative impacts experienced by host cities/countries, the number of bids for hosting both the Summer and Winter Olympics is on a declining trend. Some scholars are exploring diverse strategies to address this issue, such as establishing permanent host cities for the Summer and Winter Olympics respectively, or segmenting the Olympics into four groups based on seasonality and holding them separately. These options and strategies aim to mitigate risks borne by bidding countries and cities while also attracting greater international participation in the Olympic Games.

Figure 1: LOGO:five rings

To ensure the success of the Olympics and foster global connections through sports, ICMG suggests establishing measurement standards from various angles to evaluate the impacts of hosting the event, including land use, economic effects, human satisfaction, tourism, opportunities for future improvements, and the reputation of the host city/country. Participants are invited to make recommendations supporting ICMG's work, taking into account feasibility, implementation timelines, and the potential impact of their strategies on the set indicators[1]. Participants should also write a memorandum to the National Olympic Committee describing their proposed strategies and policy recommendations. These suggestions and policies will help ensure a successful Olympics while promoting global sports development and cultural exchange.

1.2 Our Work

The assignment requires us to develop measurement standards to evaluate the impacts of hosting the Olympics from different angles while considering the proposal's feasibility, implementation timeline[2], and the potential impact of strategies on the set indicators. Our plan is as follows:

- Establishing an evaluation model for the effectiveness of hosting previous Olympic Games.
- Based on indicators data from previous bidding countries and the impact of the first two Olympic Games, establish a willingness evaluation model.
- Testing the accuracy and sensitivity of the evaluation model.
- Providing opinions and suggestions based on the aforementioned analysis.

2 Assumptions and Justifications

To streamline our analysis, we will operate under the following basic assumptions, all of which are reasonable and valid.

● **Assumption1:**Assume that every bidding country for each Olympic event has a "strong willingness" to host.

↬ **Justification:**As the Olympics is a globally renowned event, every host country must compete vigorously with others and invest considerable resources into the event. As such, we can assume that each bidding country has a strong desire to host the Olympics, having weighed the costs and benefits carefully before submitting their bid[3].

● **Assumption2:**Assume that the willingness strength of each bidding country for each Olympic event is influenced by the success level of the host countries from the previous two events.

↬ **Justification:**Considering the massive investment of manpower and resources required to host the Olympics, it is crucial to carefully evaluate the effectiveness of past events before making any decisions about hosting future ones.

● **Assumption3:**Assuming that the Olympics are not significantly impacted by extreme and uncontrollable factors (such as war, pandemics), we consider that any changes in a host city's evaluation indicators during an Olympic year are mainly due to the impact of hosting the event.

↬ **Justification:**Due to the impact of uncontrollable factors such as war and pandemics, there have been some exceptional cases in Olympic history, such as the 1980 Moscow Olympics and the 2021 Tokyo Olympics. To keep our model more universally applicable, we will temporarily exclude these events from consideration. Furthermore, given the unprecedented scale and influence of the Olympics, as well as the high correlation between the event and our selected indicators, we will focus on major factors and attribute any changes in a host city's evaluation indicators during an Olympic year to the impact of hosting the event.

● **Assumption4:**To ensure the relevance and accuracy of our analysis, we will focus solely on data concerning the Olympic events that have taken place following World War II.

↬ **Justification:**Given the significance of World War II in shaping global history, we acknowledge that it marks an important turning point in the evolution of the modern world.[4] Therefore, to ensure our conclusions reflect the present-day context, we will focus solely on data related to the Olympic events held since the end of WWII.

The willingness to bid for hosting the Olympic Games can be influenced by the past hosting effects. Therefore, we have collected data on all previous bidding cities and established an Olympic evaluation model to score each Olympics based on a set of evaluation criteria. The data from the two previous Olympic Games before the bidding will be used as part of the evaluation index, along with several other appropriate indicators. Due to the vast number of bidding countries, we first carried out clustering analysis and then used factor analysis to calculate the weight of each indicator, ultimately obtaining a complete model for evaluating the willingness to bid. This model reflects the factors that affect willingness and their corresponding weights. After conducting accuracy tests and sensitivity analysis on the model, we will provide relevant recommendations to the International Olympic Committee (IOC).

In summary, the entire modeling process is shown below:

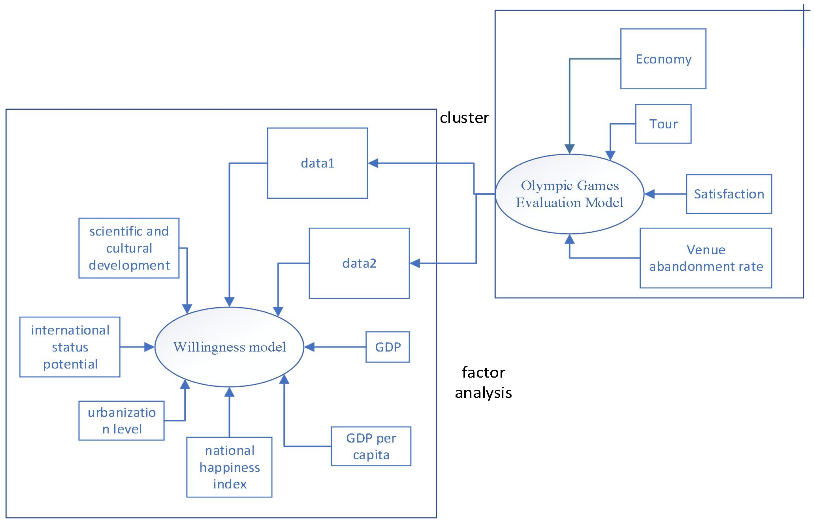


Figure 2: Schematic diagram of the modeling process

Model Preparation

This section outlines the preliminary steps necessary for modeling, including a description of the mathematical symbols used throughout the article and details on how we collected and processed the relevant data.

2.1 Notations

Table 1 displays the key symbols utilized throughout this article.

Table 1: Notations

Symbol	Description
D_i^+	Distance of the i th evaluation object from the maximum value
D_i^-	Distance of the i th evaluation object from the minimum value
$H(x)$	Information entropy
d_j	Information utility value
$V(i,j)$	Degree of willingness of the sponsoring country
$isV(i,j)$	Is the j-th country involved in the bid for the i-th
$\hat{y}_{t+1}$	Predicted values for period $t+1$
w_i	Weights of observations in period $t-i+1$
N	Number of weights
$d(x_i,x_j)$	Distance between sample points

2.2 The Data

As there are significant differences between the Winter Olympics and Summer Olympics in multiple aspects, we will treat these two events separately. Given the sizable volume of data involved, as well as its complexity and diversity, we will utilize visualizations to display some of this data more intuitively[5].

2.2.1 Data Collection

Our data collection efforts have yielded two distinct datasets, as illustrated in Table 2. The left column contains data related to each Olympic event's host city, while the right column provides information on each bidding country's characteristics.

Table 2： Resources

Data Name	Source	Data Name	Source
GDP 7 years before hosting the Olympics	https://www.imf.org/en/Home	Calendar year GDP of the host country	https://ourworldindata.org/
Number of tourists 7 years before hosting the Olympic Games	https://ourworldindata.org/	Per capita GDP of the host country in the calendar year	https://ourworldindata.org/
Population satisfaction	https://zh.wikipedia.org/	Government expenditures of the sponsoring country for the calendar year	https://ourworldindata.org/
Number of venues built	https://zh.wikipedia.org/	Population change in the host country over the years	https://ourworldindata.org/
Amount of investment	https://olympics.com/en/	Changes in investment in education over the years in the sponsoring countries	https://ourworldindata.org/
Olympic revenue	https://olympics.com/en/	Cultural influence ranking of the bidding countries	www.usnews.com
Number of abandoned venues	https://zh.wikipedia.org/		

2.2.2 Data Cleaning

Given the significant volume and variety of data collected, our work included the following steps:

- To ensure consistency and comparability between different types of data, we standardized their respective units of measurement. Specifically, we converted GDP, Olympic revenue, and investment figures to million US dollars, while tourist numbers were expressed in tens of thousands.

- For non-numeric data such as satisfaction levels, we utilized an expert system to convert these qualitative measures into quantitative data that could be analyzed alongside other numeric data points.

- To account for the potential impact of exchange rate fluctuations and inflation over the course of our extensive dataset, we have standardized all financial data relative to the purchasing power of the US dollar as of April 1, 2023, according to Eastern Standard Time.

- To ensure accurate calculations for GDP and tourist figures related to Olympic host cities, we obtained data from the six years preceding each Olympic event. Using an adaptive filtering method, we extrapolated the GDP and tourist numbers for the year of the Olympic event. We then computed the percentage difference between the actual and ideal values relative to the actual value, using this figure as a benchmark indicator for our analysis. The basic formula is as follows:

$$\hat{y}_{t+1} = w_1 y_t + w_2 y_{t-1} + \ldots + w_N y_{t-N+1} = \sum_{i=1}^{N} w_i y_{t-i+1} \tag{1}$$

Please consult the Notations section of this article for a detailed explanation of the symbols utilized throughout.

- Given the significant amount of data involved for investment amounts and Olympic revenue, we have merged these two indicators and expressed them as an investment return rate.

- As part of our analysis, we have utilized the ratio of abandoned venues to the total number of constructed venues as a key evaluation indicator, known as the venue abandonment rate.

3 MODEL I:Olympic Games Evaluation Model

Drawing on the processed data, we have identified five key evaluation indicators for constructing our model: Contribution to GDP and tourism, Venue abandonment rate, public satisfaction, and ROI[6]. We then employ a TOPSIS model that weighting by entropy method to calculate a score for each Olympic event, providing us with a measure of its overall effectiveness.

Table 3: Indicators of previous Summer Olympics

Time	GDP	Tour	Venue abandonment rate	Satisfaction&Impact	ROI
2016	0.0323	-0.0008	0/33	8	0.6267
2012	-0.0006	0.2412	0/34	8	0.8682
2008	0.0311	0.5260	1/37	10	0.7182
2004	0.0518	0.3269	2/34	8	0.7327
2000	0.0177	0.1815	2/40	8	-0.2440
1996	0.0119	0.2787	0/26	8	0.2259
1992	0.0185	0.3123	2/29	5	0.8072
1988	-0.0220	0.2059	1/31	6	-0.0256
1984	0.0000	0.0967	1/28	8	0.9450
1976	0.0149	0.1358	2/27	8	0.9900
1972	-0.0147	0.2649	2/31	4	0.9750
1968	-0.0035	0.1724	1/18	6	0.9945
1964	-0.0180	0.2759	2/19	8	0.9467

1960	0.0192	0.1979	2/17	8	0.9463
1956	-0.0277	0.2500	2/16	7	0.9663
1952	0.0209	0.4148	1/21	9	0.9866

Table 4: Indicators of previous Winter Olympics

Time	GDP	Tour	Venue abandonment rate	Satisfaction	ROI
2018	0.0494	0.1175	2/28	6	-0.3926
2014	0.2088	0.1789	2/19	6	-0.9147
2010	-0.0003	0.0338	2/14	8	-0.4365
2006	-0.0015	0.0189	2/17	8	-0.2650
2002	-0.0087	-0.0100	2/14	7	0.0455
1994	0.0036	0.4750	2/17	8	-0.8592
1992	0.0000	0.2730	0/13	9	-0.7950
1988	-0.0033	0.3588	0/7	7	-0.2400
1984	0.0025	0.1412	2/9	8	3.2300
1980	-0.0201	-0.0400	0/17	5	-0.4727
1976	-0.0033	0.1250	0/8	7	0.0250
1972	0.0201	0.0072	1/11	8	-0.6367
1968	0.0830	0.0684	0/12	8	-0.9540
1964	0.0345	0.1709	0/7	8	-0.5720
1960	0.2301	0.1630	0/10	6	-0.6000
1956	0.0255	0.2105	1/5	8	-0.8167

3.1 Model construction

The TOPSIS method is a widely utilized approach to comprehensive evaluation that effectively leverages the information contained within raw data to capture the differences between different evaluation schemes. In order to minimize the impact of subjective biases, we have employed an entropy weighting technique to determine the relative weights of each indicator used in our analysis.

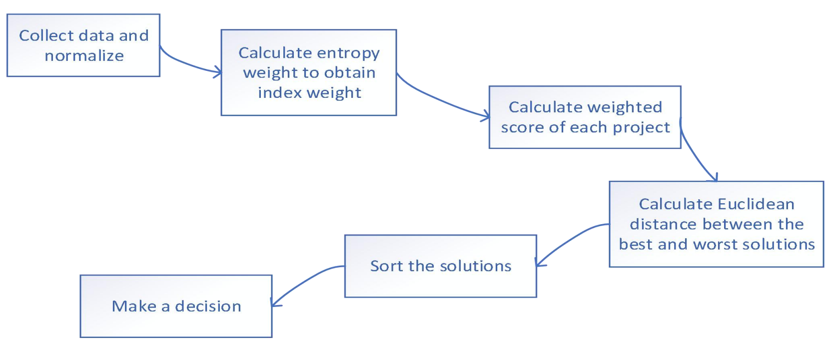


Figure 3: Flowchart of Model 1

STEP1 To begin, we further processed the data collected for each indicator and computed the rank sum ratio (RSR) for each.

$$RSR_i = \sum_{j=1}^{m} \frac{R_{ij}}{m^*n} \tag{2}$$

Here, m is the number of indicators， n is the number of groups， R_{ij} is the rank of each indicator， and the RSR value is the average rank across all indicators. The higher the RSR value, the better.

STEP2 We standardize all indicators to be maximization type indicators (the higher, the better) through a normalization process[7]. The transformation function form is not unique. Assuming the original data sequence is x， the following method is used to transform each type of indicator into a maximization type indicator.

$$X = \begin{cases} 1 - \dfrac{a-x}{M}, x < a \\ 1, a < x < b \\ 1 - \dfrac{x-b}{M}, x > b \end{cases} \tag{3}$$

STEP3 To account for potential variations in scale or units across different indicators, we normalize the positive matrix to eliminate any influence stemming from these differences. Assuming that the positive matrix is denoted as X， the resulting standardized matrix Z may be calculated as follows:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2}} \tag{4}$$

STEP4 Calculate the scores and normalize them. Assume that there are n objects to be evaluated, m evaluation indicators with sizes of nm , and a standardized matrix Z .

Define Maximum Value:$Z^+ = (z_1^+, z_2^+, ...z_m^+)$

Define Minimum Value:$Z^- = (z_1^-, z_2^-, ..., z_m^-)$

Define the Distance between the i-th Object and the Maximum Value:

$$D_i^+ = \sqrt{\sum_{j=1}^{m} (z_i^+ - z_{ij})^2}$$

Define the Distance between the i -th Object and the Minimum Value:

$$D_i^- = \sqrt{\sum_{j=1}^{m} (z_i^- - z_{ij})^2}$$

It is possible to calculate the non-normalized score of the i -th object:$S_i = \dfrac{D_i^-}{D_i^+ + D_i^-}$

Finally, we normalize the scores (normalizing the scores does not affect the ranking): $\tilde{S}_i = \frac{S_i}{\sum_{i=1}^{n} S_i}$

STEP5 The aforementioned process did not consider the relative importance of each indicator. To address this issue, we may introduce the entropy weighting method to determine the appropriate weights for each indicator. The entropy weighting method is an objective approach that assigns weights based on the principle that indicators with lower variability reflect less information and therefore correspond to lower weights (i.e., the data itself can inform us of the appropriate weighting). Specifically, the entropy weighting method assumes that events with a higher probability have lower information content, while events with a lower probability have higher information content. This relationship between information content and probability is depicted in the following graph:

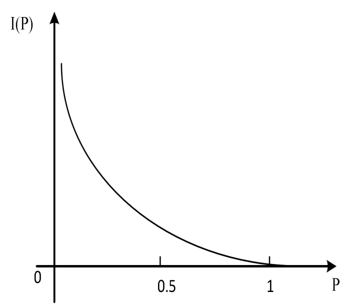


Figure 4: Entropy method information quantity as a function of probability

Consider an event denoted by x that occurs under specific conditions, with a corresponding probability of occurrence represented by $p(x)$. The entropy of this event may be defined as:

$$H(x) = \sum_{i=1}^{n} [p(x_i) I(x_i)] = -\sum_{i=1}^{n} [p(x_i) \ln(p(x_i))] \tag{5}$$

The essence of entropy is that it represents the expected value of information content. In the context of the entropy weighting method, higher entropy values indicate lower levels of information content, as they reflect a greater degree of uncertainty or variability in the data. This is because larger entropy values imply that the probability distribution of outcomes for a given indicator is more evenly spread out, indicating that the information provided by that indicator is less informative when compared to other indicators with lower entropy values. Therefore, in the entropy weighting method, we seek to assign smaller weights to indicators with higher entropy values, and larger weights to those with lower entropy values

STEP6 To calculate the weights using the entropy weighting method, we first check for the presence of negative values in the judgment matrix[8]. If any negative values are detected, we must re-standardize the matrix to the non-negative range to ensure that all elements are non-negative when computing probabilities in subsequent steps. This process can be achieved as follows:

Suppose there are n objects to be evaluated, each assessed using m evaluation indicators that have already been normalized. Further, let X denote the normalization matrix. Then, each element of the resulting standardized matrix denoted by Z satisfies

$$Z_{ij}=\frac{x_{ij}}{\sqrt{\sum_{i=1}^{n}x_{ij}^{2}}} \tag{6}$$

If there are negative numbers in the matrix Z , another normalization method should be used on :

$$\tilde{z}_{ij}=\frac{x_{ij}-\min\{x_{1j},x_{2j},...,x_{nj}\}}{\max\{x_{1j},x_{2j},...,x_{nj}\}-\min\{x_{1j},x_{2j},...,x_{nj}\}} \tag{7}$$

Calculate the weight of the j -th sample under the i -th indicator, and treat it as the probability used in the calculation of relative entropy:For the non-negative standardized matrix obtained in the previous step, calculate the probability matrix P , with each element denoted by calculated using the following formula:

$$P_{ij}=\frac{\widetilde{z_{ij}}}{\sum_{i=1}^{n}\widetilde{z_{ij}}} \tag{8}$$

Calculate the entropy of each indicator, and compute the information utility value. Normalize the resulting entropy weights to obtain the final weights for each indicator. For the j -th indicator, the formula for calculating its entropy is:

$$e_j=\frac{1}{\ln n}\sum_{i=1}^{n}p_{ij}\ln(p_{ij}) \tag{9}$$

The information utility value represents the level of information content provided by each indicator relative to the other indicators. It is defined as $d_j=1-e_j$

Normalizing the information utility values yields the entropy weight for each indicator:

$$W_j=\frac{d_j}{\sum_{j=1}^{m}d_j}\quad j=1,2,...,m \tag{10}$$

STEP7 The TOPSIS method outlined above does not account for differences in the importance of different indicators. However, when using the entropy weighting method to determine the weights of each indicator, we can simply recompute the matrices D^{+} and for each evaluated object, incorporating the appropriate weights into the calculations. The other steps involved in the TOPSIS algorithm remain the same.

$$
\begin{cases} D_i^+ = \sqrt{\sum_{j=1}^{m} w_j (Z_j^+ - z_{ij})^2} \\ D_i^- = \sqrt{\sum_{j=1}^{m} w_j (Z_J^- - z_{ij})^2} \end{cases} \tag{11}
$$

3.2 Results

Suppose that we have combined the relative errors of GDP and return on investment into a single economic indicator, and combined the relative errors of people's satisfaction, venue abandonment rate, and tourist numbers into four distinct categories of indicators. We use these indicators to score each Olympic Games based on their performance across these dimensions. By following the steps outlined above, we can derive the weights of each indicator in the scoring process. These weights will allow us to more accurately assess the overall performance of each Olympic Games, and identify areas for improvement in future events.

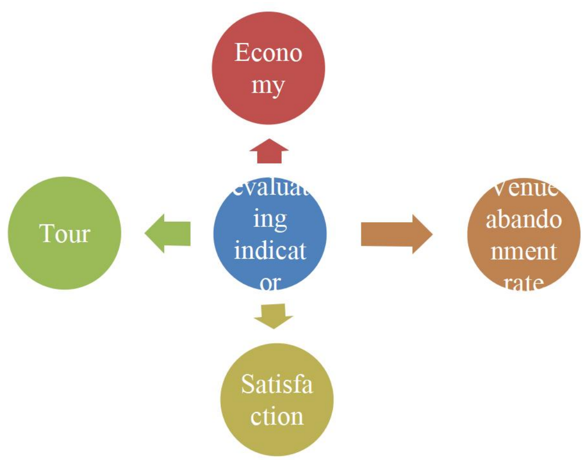


Figure 5: Evaluation metrics for Model 1

The results are shown in the following table:

Table 5: Weighting of the impact of each indicator on the Olympic Games score

Olympic	Economy	Tour	Satisfaction	Venue abandonment rate
summer	0.3021	0.2171	0.1962	0.2845
winter	0.2739	0.3855	0.1907	0.1499

Using the weights and indicators derived from the above analysis, we have computed the scores for each host country across all Olympic Games. These scores are presented in the following table:

Table 6: Scores of previous Winter Olympics

Time	Host country	Score	Time	Host country	Score
1956	Italy	0.0616	1992	France	0.0839
1960	America	0.0600	1994	Norway	0.0926
1964	Austria	0.0667	1998	Japan	0.0507

1968	France	0.0585	2002	America	0.0362
1972	Japan	0.0480	2006	Italy	0.0501
1976	Austria	0.0534	2010	Canadn	0.0488
1980	America	0.0337	2014	Russia	0.0591
1984	Yugoslavia	0.0633	2018	Korea	0.053
1988	Canadn	0.0804			

Table 7: Scores of previous Summer Olympics

Time	Host country	Score	Time	Host country	Score
1952	Finland	0.0738	1988	Korea	0.0629
1956	Australia	0.0371	1992	Spain	0.0556
1960	Italy	0.0470	1996	America	0.0812
1964	Japan	0.0474	2000	Australia	0.0812
1968	Mexico	0.0377	2004	Greece	0.0839
1972	Germany	0.0447	2008	China	0.1006
1976	Canada	0.0538	2012	England	0.0711
1984	America	0.0556	2016	Brazil	0.0662

4 MODEL II: Willingness model

To build a model for assessing a country's willingness to host the Olympic Games, we will consider several factors, including the scores obtained by previous hosts (as described in Section 4.2) and the current state of the country across various indicators. Specifically, we will focus on countries that have expressed a strong interest in hosting the Olympics in the past and use their commonalities across relevant indices to anchor our analysis.

By examining the features shared by these countries, we can create a set of criteria for determining which countries are most likely to be interested in hosting the Olympics. These criteria might include a high level of economic development, a strong infrastructure base, political stability, and a favorable public opinion towards hosting large-scale events. We will then use regression analysis to identify the relative importance of each criterion and construct a predictive model for assessing a country's willingness to host the Olympics at a given point in time.

Ultimately, this model can help stakeholders better understand the factors that drive a country's decision to host the Olympics and provide guidance for future event planning and management.

4.1 Statistics on the willingness of the applicant countries

To better understand the factors that influence a country's willingness to host the Olympic Games, we will begin by examining the bidding history for each host country across all previous Olympic Games. For each country, we will record whether they applied to host the event or not, using "1" and "0" to denote bidding and non-bidding years, respectively (represented by " $isV(i,j)$ " symbol). Additionally, if a country consistently bids for multiple consecutive Olympic Games, its willingness to host the event is likely to be higher than other countries. To account for this, we will consider the bidding history from two consecutive Olympic Games as a reference point when evaluating a country's willingness to host the event for a given year. Thus, we can calculate a country's willingness to bid for the Olympics in a given year using the following equation:

$$V(i,j) = \sqrt{\frac{\sum_{k=i-2}^{i+2} (k-i)^2 * isV(i,j)}{5}} \tag{12}$$

The results are presented in the following graph:

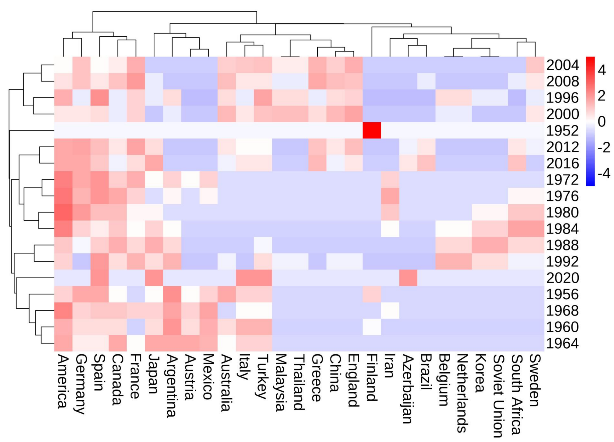


Figure 6: Heat map of the willingness of countries bidding to host the Summer Olympics

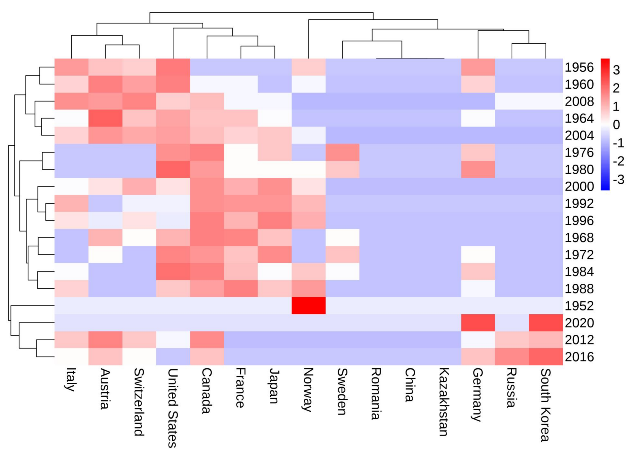


Figure 7: Heat map of the willingness of countries bidding to host theWinter Olympics

To better understand the factors that influence a country's willingness to host the Olympic Games, we will group bidding countries into 4 categories. These categories will be determined based on each country's level of economic development, infrastructure capacity, political stability, and other relevant indicators.

We will then use clustering techniques to group countries with similar characteristics together within each category. For example, developed countries may be clustered based on their GDP per capita, infrastructure quality, and level of public support for hosting large-scale events. Developing countries may be grouped based on their population size, current levels of investment in infrastructure, and degree of political stability.

Finally, we will use factor analysis to determine the relative weight of each indicator in influencing a country's willingness to bid for the Olympics, taking into account potential interactions between different variables. This analysis will help us identify the key drivers of Olympic host country willingness, and provide insights into how different factors may evolve over time in shaping a country's decision to bid for the event.

4.2 Clustering

Clustering analysis can be classified into two main categories: Q-type and R-type clustering, depending on the nature of the objects being grouped. Q-type clustering is used to group together samples that share common features or characteristics, and can provide a more nuanced and thorough understanding of the data than traditional qualitative classification methods.

Suppose each sample has p indicators, and the observed values are denoted by:

$$x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T, i = 1, 2, 3, \dots, n$$

Each sample x_i can be seen as a point in an p dimensional space. The distance between each sample can be used to measure their similarity or dissimilarity. The distances between samples must satisfy the following conditions:

$$\begin{cases} d(x_i, x_j) \geqslant 0, d(x_i, x_j) = 0 \Longleftrightarrow x_i = x_j \\ d(x_i, x_j) = d(x_j, x_i) \\ d(x_i, x_j) \leqslant d(x_i, x_k) + d(x_k, x_j) \end{cases} \tag{13}$$

When the distance does not satisfy the third condition in Equation (13), we use Euclidean distance instead:

$$d_{ij} = \left[\sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \right]^{0.5}$$

4.3 Clustering results

We have collected data on eight different indicators for all bidding countries over the past 70 years, including GDP per capita, national GDP, government expenditure, urbanization level, the proportion of exports in GDP, national happiness index, population size, and military spending as a percentage of GDP. Each year, we use clustering techniques to group all countries based on their performance across these indicators, taking into account the changing developmental status of different countries over time. Here are the clustering results for four different periods. To obtain the clustering results for a specific year, one simply needs to input the relevant indicator data for that year into the model.

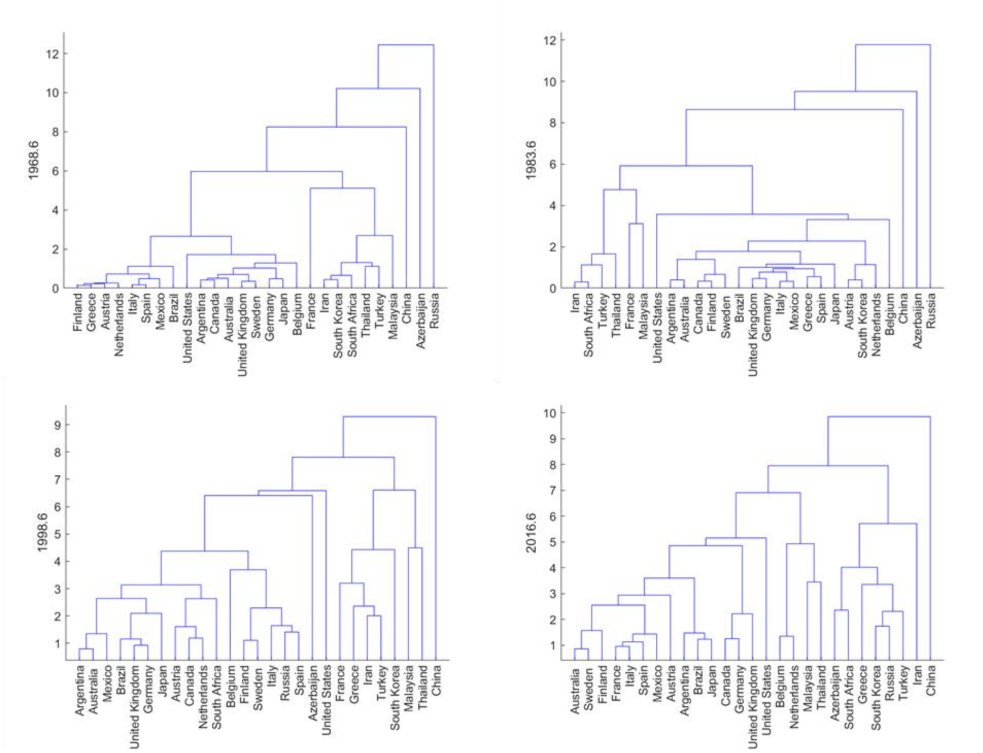


Figure 8： Summer Olympic Games bid countries classification results dendgram

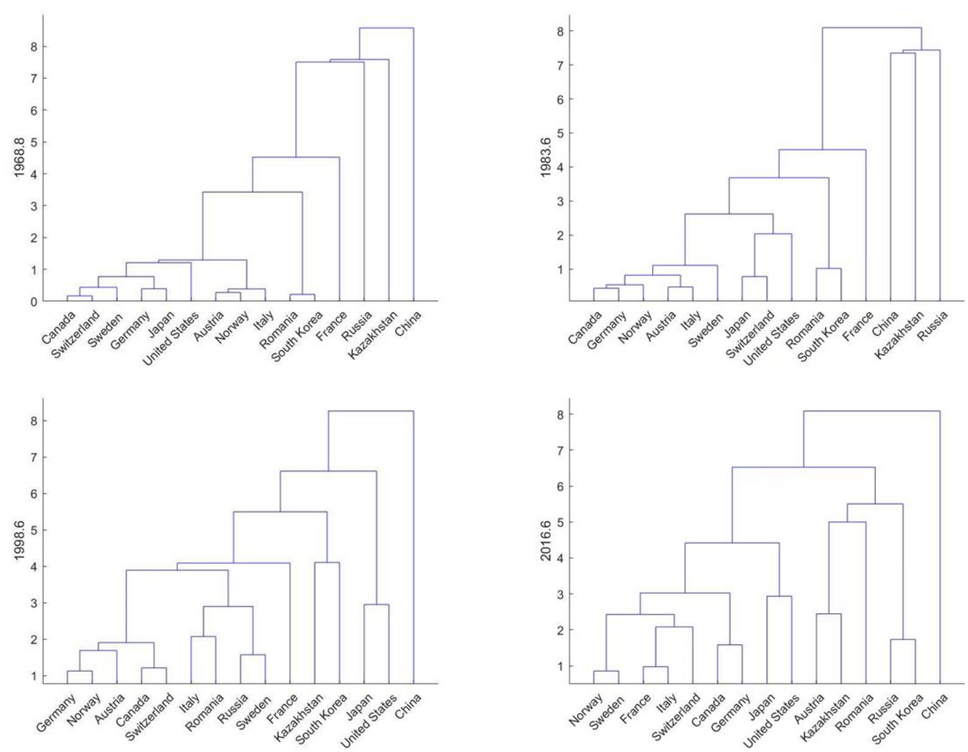
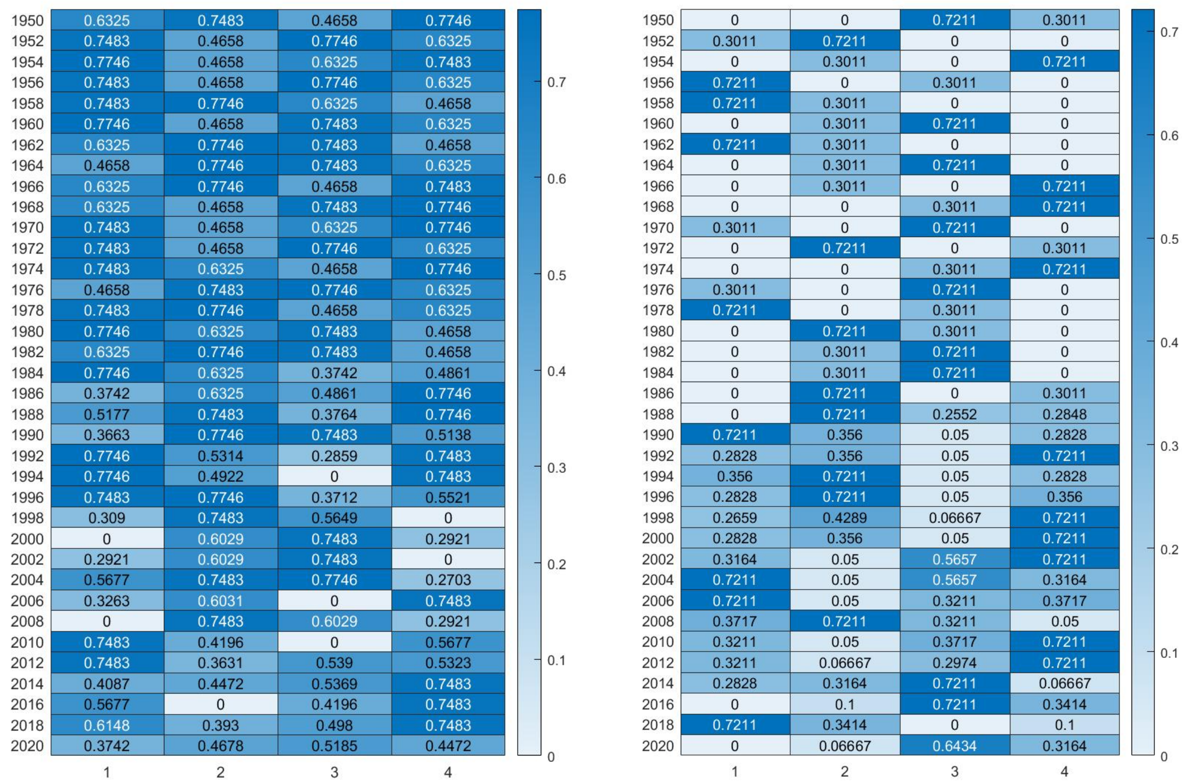


Figure 9： Winter Olympic Games bid countries classification results dendgram

The heat map of willingness to bid over time for these four categories of countries is shown below:



Summer Winter

Figure 10：Clustering results for countries willing to bid for the Summer Olympics

4.4 Factor Analysis

We have identified 4 different indicators as key factors that impact a country's willingness to host the Olympic Games. These factors include GDP growth, per capita GDP growth, the perceived success of previous Olympic Games, urbanization level, scientific and cultural development, international status potential, national happiness index. To understand the relative importance of each of these factors on the final outcome, we will use factor analysis.

Factor analysis is a statistical technique used to study the underlying structure of a set of variables by extracting common factors from them. Common factors refer to the hidden factors that are intrinsic to different variables. The basic idea behind factor analysis is to group variables into different sets based on their inter-correlations, such that variables within each group have a high degree of correlation with each other, while variables in different groups are uncorrelated or have low correlation. Each group of variables represents a basic underlying structure, i.e., a common factor.

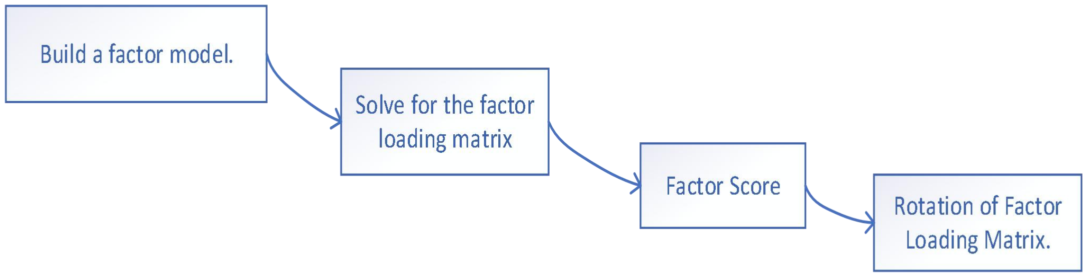


Figure 11: Factor analysis flow chart

STEP1 To ensure independence between the indicators and ease of calculation, we have categorized GDP and per capita GDP as economic factors, denoted as F_1 . The level of urbanization and government fiscal burden are considered social development factors, denoted as F_2 . The ratio of education expenditure to GDP and changes in the national happiness index are categorized as cultural and educational factors, denoted as F_3 . The hosting experiences of the previous two Olympic Games are considered an experience factor, denoted as F_4 . Additionally, the country's international status level is selected as a special factor and denoted as ε.

The common factors in factor analysis are latent variables that cannot be directly observed but have an objective existence as shared underlying factors. Each variable can be expressed as a linear function of the common factors and an error term, i.e.,

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m + \varepsilon_i, (i = 1, 2, \ldots, p) \tag{14}$$

In the above equation, $F_1, F_2, \ldots, F_m$ represents the common factors, and, X_i represents the unique or specific factors of X_i . This model can be represented by matrices as:

$$X = AF + \varepsilon$$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix}, F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_M \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

Furthermore, the model has the following characteristics:

i: $m \leqslant p$;

ii: $Cov(F, \varepsilon) = 0$, That is ,the common factors and unique or specific factors are uncorrelated with each other in factor analysis.

ii: $D_F = D(F) = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} = I_m$, That is, each common factor is uncorrelated and has a variance of 1.

iv: $D_{\varepsilon}=D(\varepsilon)\begin{bmatrix}\delta_1^2 & & & 0\\ & \delta_1^2 & & \\ & & \ddots & \\ 0 & & & \delta_p^2\end{bmatrix}$ The special factors are not correlated and the variances are not required to be equal.

The matrix in the model A is called the factor loading matrix and a_{ij} is called the factor "loading" and is the loading of the i -th variable on the j - th factor. If the variable X_i is considered as a point in the m dimensional space, then a_{ij} represents its projection on the coordinate axis.

STEP2 Solving for the factor loading matrix. We use the principal factor method to solve the factor loading matrix. The principal factor method is a modification of the principal component method, assuming that we first perform a normalized transformation of the variables. Then

$$R=AA^{'}+D$$

$$R^{*}=AA^{'}=R-D$$

Call R is the approximately correlated matrix,R The elements on the diagonal are h_i^2,not h_i, and we assume that $\hat{h}_i^2$ is an estimate of h_i^2 . Then we have:

$$R^{*}=R-\hat{D}=\begin{bmatrix}\hat{h}_1^2 & r_{12} & \cdots & r_{1p}\\ r_{21} & \hat{h}_2^2 & \ldots & r_{2p}\\ \vdots & \vdots & \vdots & \vdots\\ r_{p1} & r_{p2} & \ldots & \hat{h}_p^2\end{bmatrix} \tag{15}$$

STEP3 Rotation of the factor loading matrix. Let Q be an m order orthogonal matrix such that $B=AQ$,then:

$$BB^{T}=AA^{T}$$

Since the factor loading matrix obtained in the previous subsection is not unique, in fact the new matrix obtained after the orthogonal transformation of the matrix can be regarded as the factor loading matrix.

When we obtain an estimate of the factor loading matrix, it is possible that multiple variables have large factor loadings on the same factor or that one variable has large loadings on multiple factors. This can make it difficult to interpret or name the factors. In this case, we may want to rotate the factor loading matrix to obtain a new, simplified matrix where the factor loadings are more distinct and easier to interpret.

STEP4 Calculating factor scores. After obtaining the factor loading matrix, sometimes we want to use the common factors for other studies, such as cluster analysis or regression analysis, when we want to estimate the common factors from the original variables, i.e., to obtain the factor scores.

For the model $X=AF+\varepsilon$， if the effect of the special factor ε is not considered, we can obtain $X=AF$， but the matrix A is of $p*m$ order, and in the model we require $m\leqslant p$， usually the number of factors is much smaller than the number of variables, i.e. $m < p$, so the loading matrix is not invertible and the estimation of F cannot be obtained directly, where we use the regression method for the estimation of the common factors.

4.5 Results of Factor Analysis

For the convenience of presentation, we denote each of the eight indicators by the numbers 1-8 and derive their influence weights on willingness as shown in the following table. The results have been standardized:

Table 8: Indicator weight of summer Olympic bid country

Country category	F1	F2	F3	F4
1	0.317	0.278	0.154	0.251
2	0.345	0.262	0.162	0.231
3	0.293	0.315	0.273	0.119
4	0.251	0.252	0.33	0.167

Table 9: Indicator weight of winter Olympic bid country

Country category	F1	F2	F3	F4
1	0.159	0.213	0.226	0.402
2	0.221	0.167	0.256	0.356
3	0.322	0.277	0.12	0.281
4	0.279	0.213	0.156	0.352

5 Sensitivity analysis

5.1 Model Validation

In order to verify the accuracy of the model, we selected the bidding countries for the 2024 Summer Olympics and the 2026 Winter Olympics and several randomly selected countries that did not bid to host the games, and assessed their willingness separately. The results show that the bidding countries for the 2024 Olympic Games: France and the United States have significantly higher willingness than several other countries. 2026 Winter Olympic Games presents the same results, which supports the validity of the model.

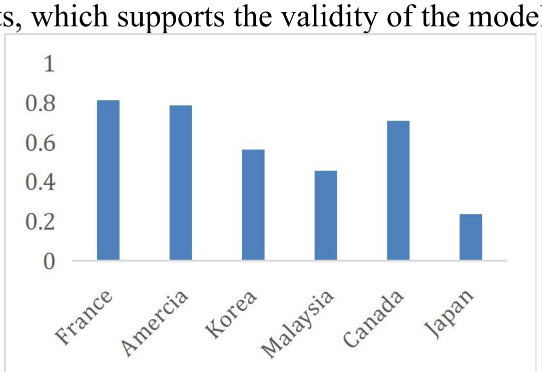


Figure 12: Summer Olympics validation results

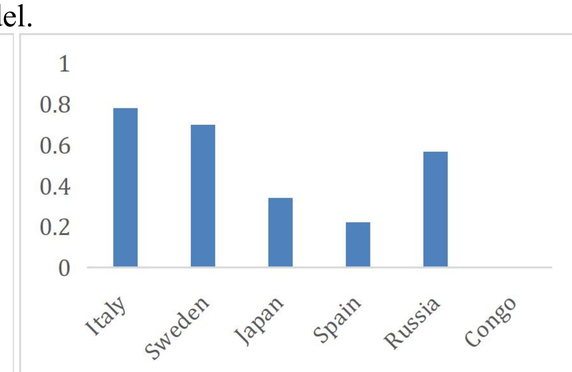


Figure 13: Winter Olympics validation results

5.2 Sensitivity Analysis

We try to do sensitivity analysis with the example of Britain's bid to host the Summer Olympics in 1996, making 4 indicators change by 10%, 20% and 50% respectively. The impact on the results corresponds to the weight of the indicator, i.e., the greater the weight the greater the impact on the results when the indicator is changed. The reasonableness of the model is proved.

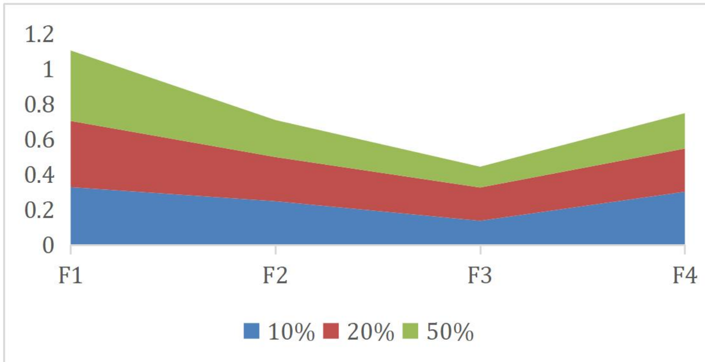


Figure 14: Sensitivity analysis results

6 Model Evaluation and Further Discussion

Based on the various data collected, we developed a model of the success of holding the Olympic Games as well as a model of willingness. The models were analyzed for sensitivity and showed good performance

6.1 Strengths

The selection of indicators used for model building is quite reasonable, taking into account multiple indicators from various perspectives that may affect the success and willingness of hosting the Olympics.

Sensitivity analysis was conducted on the model to verify its effectiveness under different circumstances.

The model was further validated for its validity by using data from countries bidding for the Olympics in the coming years.

6.2 Possible Improvements

The number of indicators selected for the model construction is limited, and the impact of the Olympics as well as the perspectives considered by countries when evaluating whether to bid for the Olympics are complex and diverse. If more indicators are selected, the conclusions will be more comprehensive.

The data collected is limited, and if more data can be obtained, the accuracy of the results will be further improved.

7 Conclusion

Based on the model developed earlier, the willingness to bid for the Olympics is influenced by a number of factors: GDP growth, per capita GDP growth, the perceived success of previous

Olympic Games, urbanization level, scientific and cultural development, international status potential, national happiness index.

The influence weight of each indicator on the degree of willingness to bid is shown in section 5.5. It is easy to see that for both the Winter and Summer Olympics, the country's GDP has the greatest degree of influence on the degree of willingness to bid, and GDP per capita and the effect of the previous two Olympic Games are equally important. Therefore, it is possible to start from these perspectives.

i: Measures to help global economic recovery

ii: Improve livelihood protection and increase GDP per capita

iii: Improving the quality of each Olympic Games, creating a positive effect and attracting more countries to participate in hosting the Games

8 Reference

[1] D. E. KNUTH The TEXbook the American Mathematical Society and Addison-Wesley Publishing Company, 1984-1986.

[2] Burgan B, Mules T 1992. Economic impact of sporting events. Annals of Tourism Research, 19 (4): 700–710. [Google Scholar]

[3] Hall Jane The Role of Chartered Physiotherapists in Olympic Sports[J] Physiotherapy, 1989, 75(12)

[4] Relationship Between Total Rewards Perceptions and Work Engagement Among Chinese Kindergarten Teachers: Organizational Identification as a Mediator [J]. Ji Dongying;Cui Li. Frontiers in Psychology. 2021

[5] Analysis of the Main Contradictions of Contemporary Sports Development and Its Solution Path[J]. Jie Yang. International Journal of Intelligent Information and Management Science. 2021(5)

[6] Wei Min, Li Shuhao. Taking the Winter Olympics as an Opportunity to Assist the Transformation and Upgrade of the Tourism Industry in Beijing and Zhangjiakou: Action Path and Guarantee Mechanism [J]. Journal of Tourism, 2020,35 (04): 7-9. DOI: 10.19765/j.cnki.1002-5006.2020.04.004

[7] Zhang Fuli Research on the Impact of the Winter Olympics on Zhangjiakou's Tourism Industry and Countermeasures [D]. Beijing Second Foreign Languages Institute, 2016

[8] Wang Kefeng, Zhang Ce, Yang Jiankun. Study on the social benefit evaluation of hosting the Beijing Olympic Games [J]. Journal of Shenyang Institute of Physical Education, 2009,28(04):28-31

Memo

To:IOC

From:Team#2330619

Date:2th April 2023

Subject: Suggestions on the hosting of the Olympic Games

Our team has conducted an in-depth analysis of various data from countries that have hosted previous Olympic Games and drawn a series of conclusions through various models. At the same time, we provide some targeted recommendations to promote the success of the Olympic Games and its implementation.

According to our findings, GDP and venue utilization are two types of indicators that have a relatively large impact on the hosting of the Olympic Games in the summer.In the Winter Olympics, economy and tourism are two categories of indicators that have a relatively large impact on the hosting of the Games.

Comprehensive above several major influencing factors, our team believes that in order to improve the enthusiasm of countries to declare the Olympic Games and promote the successful development and holding of the Olympic Games, the following aspects can be considered:

1. Provide better organization and guarantee: The organizers need to provide better and reliable guarantee in terms of venue construction, security, transportation, etc. to ensure that the Olympic Games can be held smoothly.

2. Reduce the economic burden: The organizers can reduce the economic burden of hosting the Olympics by simplifying the approval process, attracting more sponsors, and controlling the costs of preparing and holding the Olympics so that more countries can afford it.

3. Increase visibility: The organizers can increase the visibility and influence of the Olympic Games through publicity, promotion and marketing to attract more countries to participate in them.

4. Strengthen cooperation and communication: The organizers can strengthen cooperation and communication with governments, sports organizations and media to promote understanding and trust, and enhance friendship and cooperation among countries.

5. Promote sustainable development: The organizers should pay attention to environmental protection, social responsibility and sustainable economic development, encourage countries to share their experiences and practices, promote sustainable development, and increase the enthusiasm of countries to declare the Olympic Games.

Our team is confident in the above recommendations. First, all our data are based on previous Olympic Games, and all our analysis is targeted to solve the problem of declining enthusiasm to bid for the Olympic Games. Secondly, the model we built can fully explore the relationship and influence degree of each indicator with high accuracy. Finally, we have conducted model validation, which is very consistent with the reality and highly reliable.

We sincerely hope that our analysis results and suggestions can help you. Thank you for your support and trust.