Team Control Number

14977

Problem Chosen

A

2024
HiMCM
Summary Sheet

Evaluating Olympic Disciplines: A Data-Driven Integrated EWM-TOPSIS Approach

We aim to create a comprehensive and objective model that evaluates sports against all IOC criteria using a data-driven approach, improving transparency and consistency in decision-making. For all data, we evaluate by disciplines. We collected extensive data beyond Olympic statistics, to gain holistic results, ranging from Gender Ratio statistics based on previous studies to Global Interest provided by Google Trends, for all 52 disciplines. Data for every factor was normalized independently, to ensure that the score of every factor falls within the range of 0 to 1, which simplifies subsequent calculations.

We constructed an assessment model based on Entropy Weight Method (EWM) and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) to evaluate the fitness in regards of the IOC Criteria. Criteria that have more two factors had EWM and TOPSIS performed to the two factors before finding the ideal solution for the six major IOC criteria. Weights of criteria are determined through the value dispersion by entropy. With TOPSIS, we found the ideal solutions and calculated each discipline's distance from the ideal solutions. All steps for the model were iterated for 34 countries, to generate 34 sets of location-dependent discipline ranking.

Then, we tested our model to 6 disciplines that did not take part in building the model, by inserting their data to TOPSIS weighted criteria and yielding final scores. These final scores were compared against the ranking of the untested disciplines. The rankings of Breaking, Baseball and Softball, and Flagged football matched the 2028 Los Angeles Olympics' decision for these disciplines. For Rowing, Fencing, and Artistic Gymnastic that have continued to be in the Olympic since 1988, our results for 2016 to 2032 Olympics all validated their continued presence. In addition, we suggested that e-sports, cricket and frisbee, ranking 1st, 16th and 36th for Australia, may be added to the 2032 Brisbane Olympics.

Finally, we discussed the robustness of our model through sensitivity analysis by Root Mean Square Error. Out of 52 total disciplines, we randomly selected 100 sets of 43 disciplines as that determine the weights. Our model averaged an RMSE of 0.01031, demonstrating a high robustness. In fact, adding ±0.05 to every single standardized value, the model's outputs only had a RMSE of 0.01801, and thus our model provided holistic view on disciplines evaluation.

Keywords: Olympic, Entropy Weight Method, Technique for Order of Preference by Similarity to Ideal Solution, Assessment Model

Evaluating Olympic Disciplines:

A Data-Driven Integrated EWM-TOPSIS Approach

Goal

Create an **assessment model** that evaluates all IOC criteria to provide a more holistic, **objective** and **data-driven** approach in evaluating potential SDEs.

Data

Extensive data collection beyond Olympic statistics, ranging from gender ratio statistics based on previous studies to global interest provided by **Google Trends**. Take **z-score** and **normalize** all data to fall within [0,1].

Results

Model produces 34 sets of ranking based on countries, for all 49 disciplines currently in the Olympic and 3 other suggested disciplines.

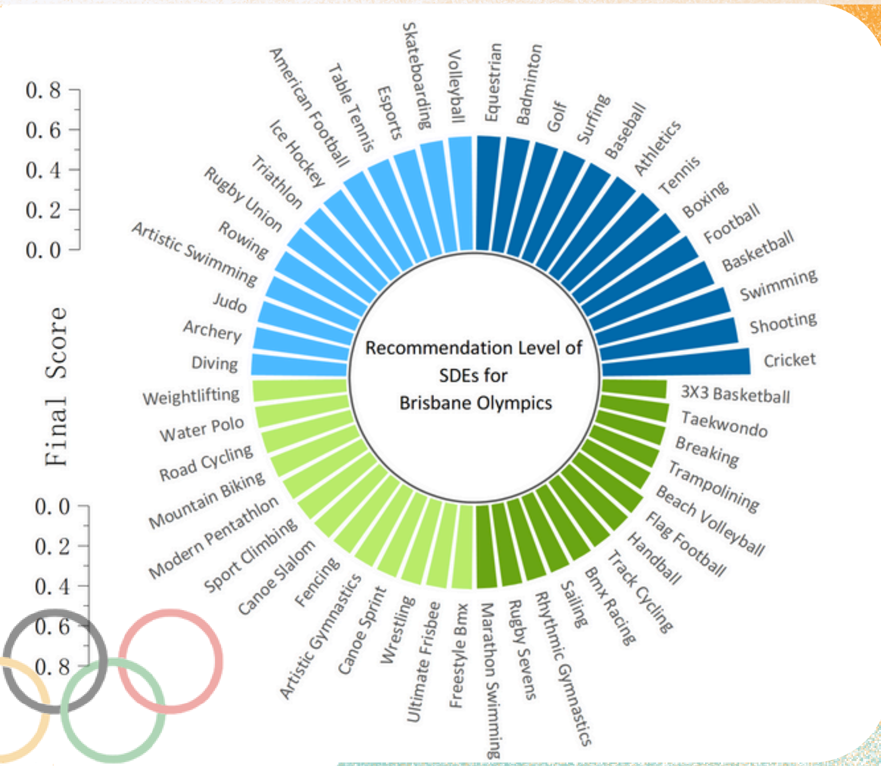

Fig. 1 Final Evaluation Score for 52 Disciplines for 2032 Brisbane Olympic

Disciplines ranked above 40 are recommended. We suggest the addition of, in order of recommendation: **cricket, e-sports and frisbee**.

Model

Uses multi-criteria decision-making approach. Model incorporates **EWM** with **TOPSIS**, to find the weight and ideal solutions for both suggested factors and IOC Criteria, based on data from 43 untested disciplines. Location-dependent results are **iterated** for **34 countries**. Disciplines' distance to the positive and negative ideals calculated and ranked.

Table of Contents

1 Introduction **4**
- 1.1 Background 4
- 1.2 Breaking Down Our Tasks 4
- 1.3 Overview of Our Work 5

2 Assumptions and Justifications **6**

3 Factors **7**

4 Data and Variables **9**
- 4.1 Data Collection 10
- 4.2 Data Preprocessing 11
  - 4.2.1 Transformation 11
  - 4.2.2 Normalization 12
  - 4.2.3 Presentation 13

5 Model **13**
- 5.1 Sub Criterion 13
  - 5.1.1 Entropy Weight Method I 13
  - 5.1.2 TOPSIS I 14
- 5.2 IOC Criterion 15
  - 5.2.1 Entropy Weight Method II 15
  - 5.2.2 TOPSIS II 15

6 Model Implementation **17**
- 6.1 Testing Against SDEs 18
- 6.2 Proposing New SDEs 19

7 Modified Model **20**
- 7.1 Sensitivity Analysis 20
  - 7.1.1 Quantifying Change in Evaluation 20
  - 7.1.2 Changing Disciplines 20
  - 7.1.3 Changing Inputs 21
- 7.2 Extensions 21

8 Conclusion **22**

9 Letter **23**

References **24**

Appendix **25**

1 Introduction

1.1 Background

Athletes aim for "Faster, Higher Stronger, - Together;" the International Olympic Committee (IOC) aims to promote global and fair Olympic Games. Since Athens 1896, having had worldwide influence for more than a century, the Olympics have been hosted once every four years [1]. Over the years, the IOC widens the scope of the game and regularly modifies the sports, disciplines, or events (SDEs) in the Olympic.

The Olympic Programme Commission of IOC develops a set of six criteria, including: Popularity and Accessibility, Gender Equity, Sustainability, Inclusivity, Relevance and Innovation, and Safety and Fair Play. However, these criteria are general and concept-based, making them difficult to be used directly for evaluation, since subjectivity and ambiguity are unavoidable. Therefore, combining both qualitative and quantitative evaluation with specified strands is highly required for the application of these criteria.

There have been attempts to evaluate the Olympic Games on individual aspect. For example, sustainability was evaluated for each Olympic Game based on scoring nine indicators and was presented and compared parallelly [2]. But the IOC's criteria are diverse, which requires considering the weight of the indicators. In addition, the individual SDEs and all six of the IOC criteria are hardly discussed.

To address this issue, we aim to create an assessment model that evaluates all IOC criteria to provide a more holistic, objective and data-driven approach in decision-making, improving the transparency and consistency of evaluating potential SDEs [3]. Extensive data preprocessing enhances the model performance and improves compatibility of data from diverse sources. By incorporating Entropy Weight Method (EWM) with Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), the model presents an objective ideal condition and compares each discipline with it. This model also provides ranking of all disciplines in the Olympic for 34 different hosting countries that allows an intuitive understanding of the position, based on the IOC criteria.

1.2 Breaking Down Our Tasks

This study completes the following tasks:

- Task 1

We propose several factors under each IOC Olympic criterion that meet the requirement of specific and measurable, such as global participation rates and athletes' gender ration. After collecting data of current disciplines, we positively transform and normalize them.

- Task 2

We build a model that can evaluate the fitness of an SDE with the Olympic criteria, dependent to the hosting country. Our model is based on the TOPSIS method. It ranks the alternatives, in this case the SDEs, based on their distance from an ideal solution. We begin with using entropy weights to determine the importance of different sub criterion — the factors we proposed in Task 1. With TOPSIS, we find the ideal solution of each criterion, with two or more factors. Then, we apply the same two steps to the IOC criterion and receive the optimal score of a discipline. Our final result is presented as 34 lists of ranking for all 49 existing disciplines and the 3 recommended disciplines, for each of the selected countries.

- Task 3

We test our Olympic Sports Evaluation model against SDEs that were recently added or removed from games or have continuously been in the Olympic since 1988. We selected the SDEs with varying levels of popularity, cost, and environmental impact. We choose: Breaking, Baseball and Softball, Flagged Football, Rowing, Fencing, and Artistic Gymnastics.

- Task 4

We suggest three SDEs that have the potential to be added to the 2032 Olympics in Brisbane. The SDEs are, in order of recommendation: cricket, e-sports and frisbee.

- Task 5

We perform sensitivity analysis using Root Mean Squared Error. By changing the disciplines for weight calculation, we find that the model is resilient to changes in disciplines used to calculate weights.

1.3 Overview of Our Work

Our work (Figure 1) is presented from Section 2 to 7. This paper is organized as followed: Section 2 states the assumptions suggested in the study; Section 3 presents the factors to specify each criterion; Section 4 demonstrates the complete process of data collection and preprocessing; Section 5 provides the detailed description of the model; Section 6 tests the model against added or removed SDEs and suggests SDEs for potential additions; Section 7 presents a sensitivity analysis on the and addresses its limitations; Section 8 concludes the study and suggests further implications.

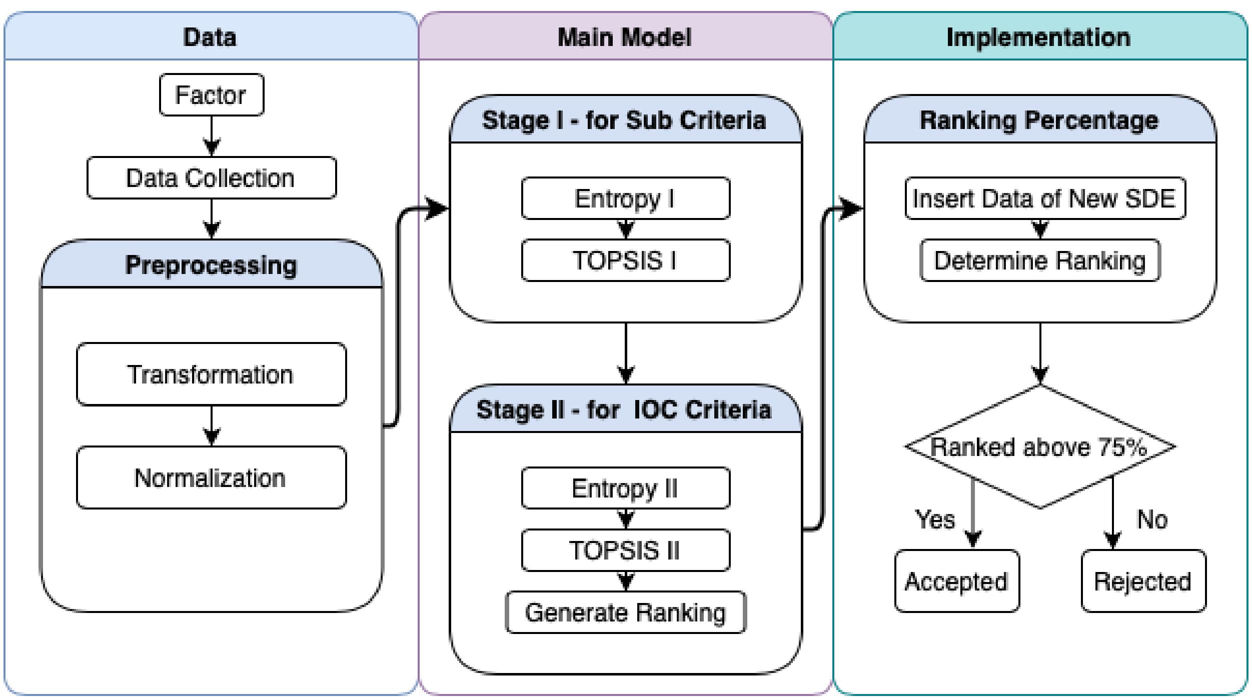

Figure 1: Flowchart Showing the General Logic of Our Work

2 Assumptions and Justifications

To clarify the problem, we propose a series of assumptions:

Assumption 1: From 2024 to 2032, the data for each sub criterion, except for the ones under popularity, will not have significant change due to factors other than the change in host country.
Justification: Major external factors, such as change in economy and climate over time, have a relatively uniform impact across all potential host countries. Therefore, their influence would not create significant variations between the sub criteria. However, popularity is strongly correlated with time, due to different trending values. Therefore, for the criterion of Popularity and Accessibility, we require recent data to mitigate the discrepancy between the popularity when the data is collected and that in the future.

Assumption 2: The Olympic data only reflects the situation of existing Olympic events and is irrelevant to SDEs not in the Olympic.
Justification: To enable our model to evaluate both current and potential Olympic events, the general public data is used instead of relying solely on Olympic data. This broader dataset provides information on criteria such as publicity and sustainability for SDEs, regardless of their current Olympic status.

Assumption 3: The environmental and social outcomes of a country's preparation to bid for the Olympic, are not considered as part of the IOC criteria Sustainability and of any factors. However, planning for and building new stadiums that may start before the country b
Justification: The IOC criteria mainly focus on the games themselves. A country's pre-bid activities, such as initiating community and building roads, are considered independent from the actual games. Instead of directly influencing the Olympic games, they are more of improving the country's own well-being and appeal.

Assumption 4: The IOC criteria are consistent over time.
Justification: Although there may be refinements and modification of the criteria in the future, in order to match with values of that time, these potential for changes will not be taken into consideration because they are unpredictable. The overall framework of the criteria is relatively stable. However, we will address the changes in weights of each criterion and factor, longitudinally.

Assumption 5: There will not be drastic social or environmental changes globally, such as World War II in 1944 and COVID-19 in 2020, that will cause the Olympic to be canceled or postponed.
Justification: Natural disasters or political changes will cause unpredictable consequences to the performance of the SDEs on each criterion. And these events have low probability to occur [].

Assumption 6: SDEs and Olympic data are accurately and comprehensively reported by bidding countries for all years.
Justification: Data availability and reliability may vary for countries. However, we trust each country to provide accurate data, especially in recent years when technology enhances communication and measurements. Measures of uncertainty will not be considered for the historical Olympics (i.e. prior to the 1950s) because inconsistency cannot be tracked.

Assumption 7: All data used to build the evaluation model does not include SDE data that are yet to be tested. The impact of the evaluated SDEs on the weights is negligible.
Justification: The model evaluates whether SDEs can be added to the Olympic, so it is centered around the Olympic. The SDEs tested are currently not in the Olympic, so their impact can be neglected.

3 Factors

For each of the following IOC criterion, we suggest one to two factors, or sub criterion, based on the definition and implication of the IOC criteria. Criteria such as Gender Equity is rather straight forward, so it will be represented through only one quantitative factor. All our factors are quantitative or have proposed quantification methods because they will need to be fitted into an evaluation model. For Sustainability, we have included 10 deterministic questions that address essential yes or no (i.e. 0 or 1) elements, a method inspired by multi criteria decision making approaches [4][5]. We also referred to some ideal of sustainable development of Olympic sports [6][7]. For Relevance and Innovation, the 10 questions for each topic were designed to cover a wide range of aspects [8]. The impact of technology on sporting performance is also considered when designing our deterministic questions [9]. To mitigate the risk of oversimplification, as deterministic factor's significant impact on the model may induce invalid results, we designed the questions to address the greatest range of topics in one criterion. The relationship between factors and criteria are shown in Figure 2.

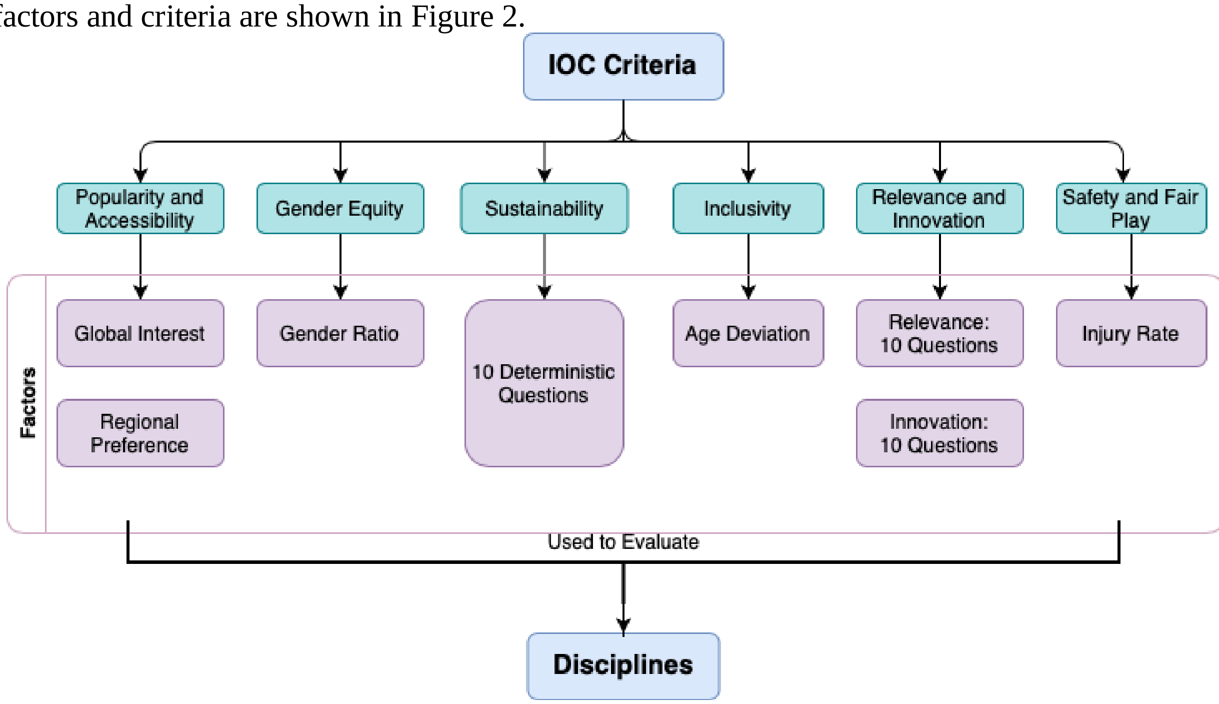

Figure 2: IOC Criteria's Relationship with Factors

I. Popularity and Accessibility

- Global Interest

Measured through the online discussion rate for each discipline provided Google trends [10]. A final score is provided for each discipline.

- Host Country's Regional Preference

The initial choice of the zero point for the variable under analysis is arbitrary, other values are all relative to its popularity inside the country. These values cannot serve as a comparison in different countries; they only serve as a relative value inside each single country. Since the final values are standardized into z-scores, the selection of the zero point does not affect the analysis or interpretation of results. Z-scores are not influenced by shifts, ensuring that all values are expressed in relative terms. This approach allows for comparability and eliminates the influence of the initial scale origin.

II. Gender Equity

- Gender Ratio

The ratio between male and female athletes in each discipline [11]. Provided by an investigation in 2009 [12], data is collected from 15859 participants recruited up to November 2005. The original data only contains gender ratio, injury rate and BMI of 17 activities. However, our task is to rank 43 disciplines and 9 disciplines tested in Section 6, and we require gender ratio for all 43+9 disciplines. Due to data limitations in the article, we manually categorized these 43+9 events into 7 major categories: Water Sports, Team Ball Sports, Individual Ball Sports, Cycling. Martial Arts and Dance, Athletics, and Board and Boat Sports. The 17 activities provided in the original data were also classified with same categories. We then used the average gender ratio for each major category as the data for all disciplines within that category.

III. Sustainability

Example deterministic questions are provided below:

1. Does the SDE require significant land alteration (e.g. deforestation, habitat destruction) for its infrastructure?
2. Are there insufficient efforts or guidelines in place to minimize the environmental impact of the activity?

Remaining questions are presented in Appendix 1.

IV. Inclusivity

- Standard Deviation in Age

The standard deviation in age of people participating in the discipline, according to a 2009 study in Spain [12]. Even though the study was done 15 years before, the result of inclusivity was not impacted by time, as stated in Assumption 1.

V. Relevance and Innovation

- Relevance

Example deterministic questions are provided below:

1. Does the sport incorporate modern technologies, such as virtual reality, augmented reality, or advanced analytics, in its gameplay or viewing experience?
2. Are there opportunities for mixed-gender or co-ed participation within the sport?

- Innovation

Example deterministic questions are provided below:

1. Does the sport or activity utilize newly developed equipment or technology in its gameplay or performance? (e.g. advanced materials, wearable sensors)

2. Has the sport or activity introduced a new way of measuring or scoring performance in recent years?

The remaining questions for both Relevance and Innovation are presented in Appendix 2.

VI. Safety and Fair Play

- Injury Rate

The risk of injury when participating in the discipline. Data is collected from the same source and in the same way as Gender Ratio [12].

4 Data and Variables

For all data, we choose to evaluate by disciplines. People talk about disciplines instead of events, and the decision of events are made after the discipline is allowed. Sports are too broad, and some sports have disciplines under it, while others do not. This will lead to discrepancy when analyzing current Olympic data.

As shown in Figure 3, to collect and process data for the sub criteria, they are first categorized based on whether they are deterministic or not. The questions under Sustainability, Relevance and Innovation are deterministic. Their data are directly evaluated, since most questions are objective and straightforward, and the answers are stated in the discipline's definition. They are then summed to directly yield a score. The weights between each question are not considered because scores for discipline are similar and many scored 0 on the same question, as presented in Section 4.2.3. If the weights are calculated through entropy, then one question will take up the significant weight, as the deviation is greatest, which causes oversimplification for that criterion. The scores of each discipline on Sustainability and Relevance and Innovation are presented in Appendix 1 and 2.

As for non-deterministic sub criteria, data that can be directly obtained will go through Section 4.2.2. The data sources are previously provided in Section 3. However, Global Interest and Host Country's Regional Preference requires special pre-processing, since data comes from Google Trends. Section 4.1 and 4.2 will focus on how data are obtained for these two factors, by using Global Interest as an example.

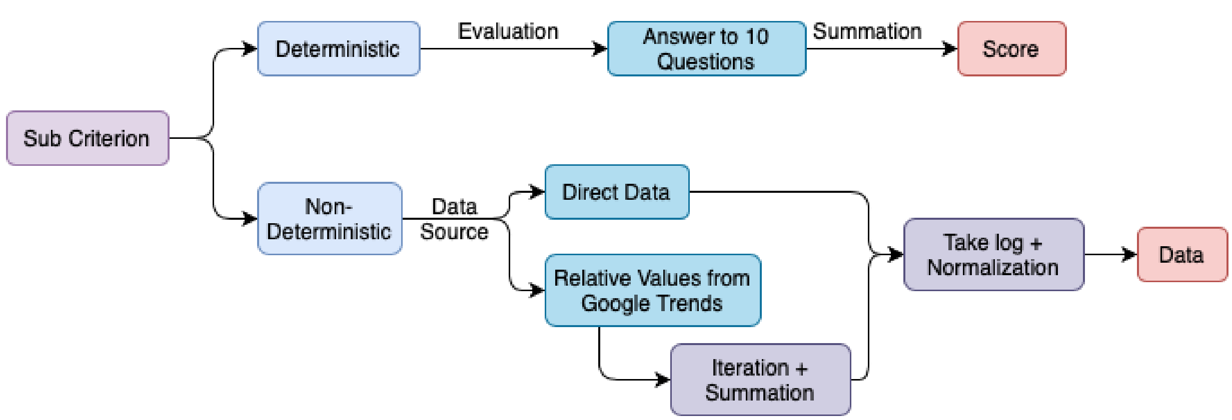

Figure 3: Logic of Data Collection and Preprocessing

4.1 Data Collection

Assumption 1 states factors are not impacted by time, except for Popularity and Accessibility. For this criterion, we consider a time span from November 12, 2019, to November 12, 2024, or the recent five years.

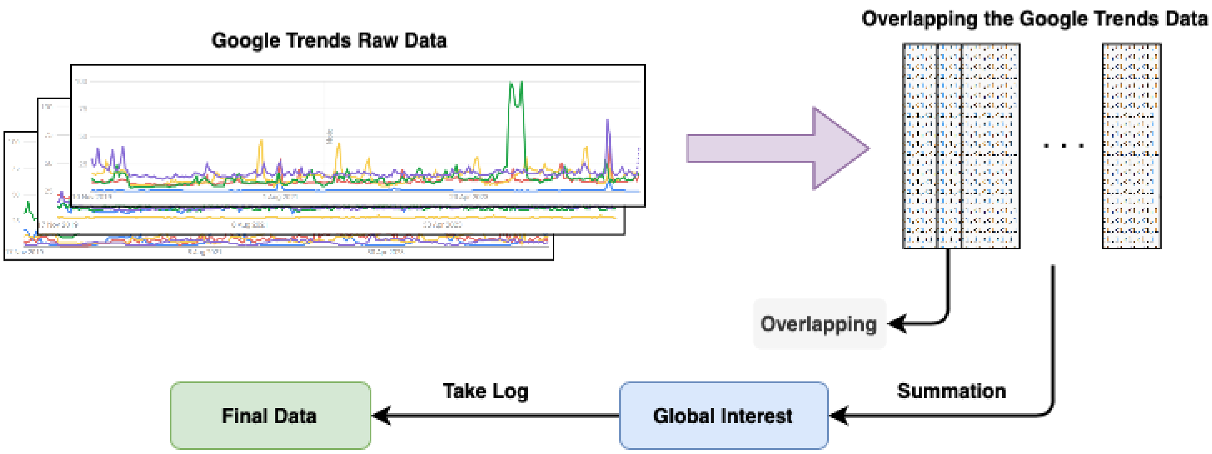

Figure 4. Flowchart of Data Collection Process from Google Trends

Our data for Global Interest is provided by Google Trends – a website that analyses the popularity of searches in Google Search across various regions [13]. The procedure of collection and preprocessing are shown in Figure 4. Based on popularity of searches on 52 different disciplines on a weekly basis, Google Trends generates a Popularity and Geographic Distribution graph. For each set of 5 disciplines over 5 years, there is one item that has a value of 100, while the values of others are relative to this item. There is a global popularity for each discipline over time, and a local popularity across each different discipline. Both are treated separately.

However, Google Trends can only compare up to five disciplines at the same time, and it provides the relative value, as shown in Figure 5 with y-axis as Relative Global Interest. To avoid the discrepancy between sets of five disciplines due to them being relative data, we assign each set to have one discipline overlapping with the previous and next set. We receive a total of eleven sets of relative values between disciplines, for each country.

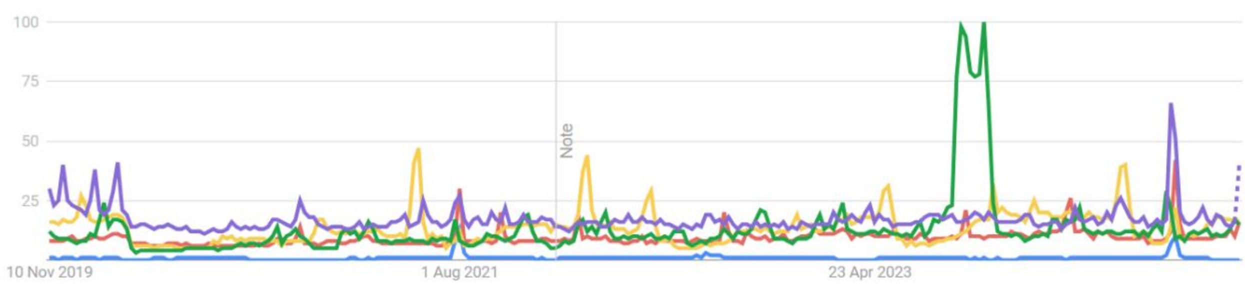

Figure 5. Example data from Google Trends

4.2 Data Preprocessing

To calculate the proportions across discipline, we combine the relative popularity into a single table. Iterating the process, we receive the weekly relative popularity from the recent five years. We arrange the disciplines in descending order based on their summed weekly values. The ranked disciplines are then assigned to new groups of five, also with overlapping items, with order.

Subsequently, the new eleven sets of disciplines are inserted to Google Trends, for a new relative value. This step is done to minimize errors. Since Google Trends only yield whole numbers for relative value, the relative value between random disciplines is randomly assigned groups may be significantly large, which will greatly impact the model, increasing the error rate within the data. For example, 0:0:1:55:100 does not provide enough information about the first three disciplines. Only when comparing disciplines with similar Global Interest will the relative value accurately present the discrepancy, such as 10:13:17:30:14. The data up to this step is shown in Appendix 3. This step changes the relative ranking of the disciplines. After multiple iterations, the rankings of the disciplines turn stable, and then we can finally utilize the data.

The values are designated to be relative to Marathon swimming, the sport with a relatively low popularity in all of the iterations. This will make the popularity be more consistent when calculated across different sets of data, since all the ratios are integers. This decision does not affect the model's final outputs, since we will take the z-score of the data later on.

4.2.1 Transformation

We realize the results contain values depicting "less than one" and zeros. Since "less than one" is not a concrete value, it cannot be used for future calculation. We tally the frequency each number appeared in the relative values. The value of "less than one" is considered according to Benford's Law, where data appears to be in a Benford distribution. By applying a logarithmic fit, we calculate the best value for it is 0.45.

To avoid having zeros when calculating ratios, we increment all values by one. Combination and summation of data for every week of 5 years are then done to the new values. Figure 6 shows the Global Interest for different disciplines in ascending order.

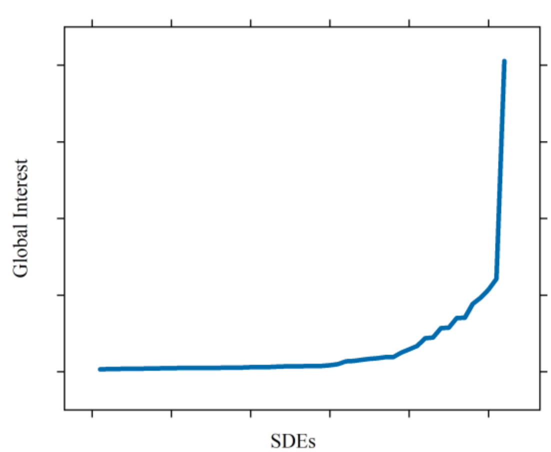

Figure 6. Global Interest Calculated through Google Trends.

4.2.2 Normalization

Google Trends Data

Data on Global Interest and Host Country's Regional Preference are taken logarithms. Prior to taking log, the data forms a power-law distribution, according to Figure 6. Taking a logarithm to the power-law distribution data makes the distribution more symmetrical and allows the usage of linear statistical methods. After the logarithm, the data is shown in Figure 7.

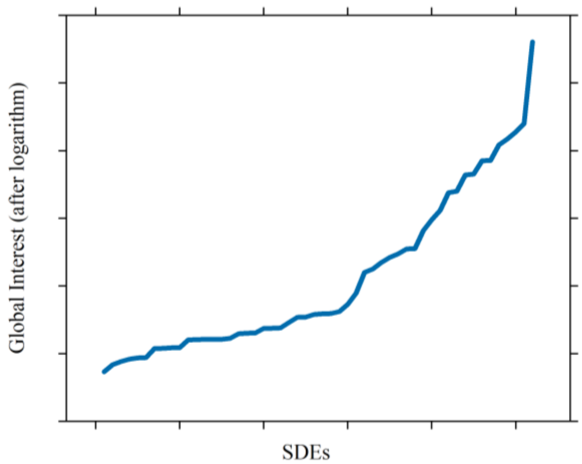

Figure 7. Global Interest After Logarithm

After taking the logarithm, we calculate the z-score of all of the current Olympic sports and use the mean and standard deviation to calculate the scores for other sports. The z-scores will be standardized to fit within 0 to 1.

Non-Deterministic Factors

Non-Deterministic Factors are normalized through slightly different methods due to the difference in their range. This approach ensures that the score of every factor falls within the range of [0, 1] and is a "higher-is-better" indicator, which simplifies subsequent calculations.

For Category 1 (Popularity and Accessibility), after calculating the z-score for both variables, we apply min-max normalization:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Injury Rate is normalized using:

$$Injury\ Rate\ Score = 1 - \frac{Injury\ Rate}{100}$$

This ensures all values fall within the range [0, 1] and represents the non-injury rate per 10 individuals.

The normalization formula for Gender Ratio is:

$$Gender\ Ratio\ Score = 1 - \left|\frac{Male\ Participants - Female\ Participants}{Male\ Participants + Female\ Participants}\right|$$

Standard Deviation of Age is normalized by dividing the value by the maximum value in the dataset, also known as max-normalization. Dividing each value by the maximum ensures that all values are scaled between 0 and 1. Also, the Standard Deviation of Age dataset contains only non-negative values, the transformation is naturally intuitive since the highest value is scaled to 1, and smaller values are proportionally scaled down.

Deterministic Factors

Data for deterministic factors are normalized by dividing each discipline's score by ten, scaling the values to between 0 and 1.

4.2.3 Presentation

The final normalized data can be found in the Appendix, with Table 1 showing the specific location. Data for Gender Ratio, Standard Deviation in Age, and Injury Rate are presented in one Excel Sheet. This data will be directly used in Section 5.

IOC Criterion	Factor	Data
Popularity and Accessibility	Global Interest	Appendix 4
	Host Country’s Regional Preference	Appendix 5
Sustainability	10 Questions	Appendix 1
Gender Equity	Gender Ratio	Appendix 6
Relevance and Innovation	10 Questions for Relevance	Appendix 2
	10 Questions for Innovation	Appendix 2
Inclusivity	Standard Deviation in Age	Appendix 6
Safety and Fair Play	Injury Rate	Appendix 6

Table 1. Data for Each Factor and Its Corresponding Location in the Appendix

5 Model

Our model is separated into two sections, each incorporated EWM with TOPSIS, which presents an objective ideal condition and compares each discipline with it, deducing an evaluation score. Criteria including Popularity and Accessibility, and Relevance and Innovation that have two sub criteria, require to perform EWM and TOPSIS to the two sub criteria before finding the ideal solution for the six major IOC criteria. Since data on Popularity and Accessibility change by locations, the model is not universal to all countries. All steps for the model are iterated for 34 countries, provided in Appendix 5. Adding and removing decisions [14] are made in Section 6. The entire model uses one complete code that is in Appendix 7.

5.1 Sub Criterion

5.1.1 Entropy Weight Method I

The EWM measures the value dispersion in decision making. We chose EWM because it can successfully balance the weights, giving the ones with greater standard deviation larger weights, even with limited number of factors [15].

For sub criterion, we begin with a matrix with three columns showing the name of the 43 discipline and its scores for the two sub criteria, respectively, which means j ranges from 1 to 2 and i ranges from 1 to 43. But for it to be reused in Section 5.2.1, we suppose we have m evaluation target, and n factors. The initial matrix of data will be:

$$X = (x_{ij})_{n\times m} \qquad (1)$$
$$i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m.$$

Based on the definition of information entropy, entropy value e_j of the jth factor is

$$e_j = -k \sum_{i=1}^{n} p_{ij} \ln p_{ij}\ ,\ with\ k = \frac{1}{\ln n} \qquad (2)$$

In the formula above, p_{ij} is the weight of the ith evaluation target, in this case the discipline, of the jth factor. It is:

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^{n} y_{ij}}$$

Then, we find out the entropy weight, which is the weight assigned to the sub criterion based on its entropy value. The entropy weight of the jth factor is calculated through:

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{m}(1 - e_j)}, \qquad with \sum_{j=1}^{m} w_j = 1$$

The score is then calculated through TOPSIS.

5.1.2 TOPSIS I

TOPSIS method allows us to find the positive ideal solution that represents the best possible outcome across all considered criteria, and the negative ideal solution that shows the least desirable outcome [16]. TOPSIS consists of two parts: finding the ideal solutions and calculating the evaluation target's distance from the ideal solutions. For the complete version of TOPSIS, the difference in conditions for positive and negative indicators need to be considered. Since the factors we chose in Section 3 are all positive indicators, we will only include the condition of positive indicators.

With j being the sub criterion and the i being the discipline, we assume the positive ideal solution X^+ is:

$$X^+ = (x_1^+, x_2^+, \cdots, x_m^+), \qquad with\ x_j^+ = max_i(x_{ij})$$

For negative ideal solution Z^-, we assume it is:

$$X^- = (x_1^-, x_2^-, \cdots, x_m^-), \qquad with\ x_j^- = min_i(x_{ij})$$

The distance between the target discipline's score to the positive ideal solution d_i^+ is:

$$d_i^+ = \sqrt{\sum_{i=1}^{n}((x_{ij} - x_j^+)^2 \times w_j)}$$

Similarly, the distance to the negative ideal solution d_i^- is:

$$d_i^- = \sqrt{\sum_{i=1}^{n}((x_{ij} - x_j^-)^2 \times w_j)}$$

We define a score C_i $(0 \leq C_i \leq 1)$ for each discipline to be:

$$C_i = \frac{d_i^-}{d_i^- + d_i^+}$$

Inserting the data to our TOPSIS model will yield a matrix with seven columns: the discipline and corresponding scores C_i for six IOC Criteria. The score C_i for all current Olympic disciplines is calculated. For criterion other than C1 (Popularity and Accessibility) and C5 (Relevance and Innovation), the score C_i will be the normalized data in Section 4.2.3. This matrix X, with 43 evaluation target, and 6 criterion will be used as initial data for Section 5.2.1. This matrix is presented in Appendix 11.

5.2 IOC Criterion

5.2.1 Entropy Weight Method II

The entropy weight of each IOC Criteria will be calculated similar to the method in Section 5.1.1. However, the index j will be the criterion, which ranges from 1 to 6. Index i is still the discipline that ranges from 1 to 43. The initial matrix in $Formula\ (1)$ was created in Section 5.2.2. And the entropy weight of each criterion is calculated through $Formula\ (2)$. An example of teach criterion's weight for Japan is presented in Figure 8.

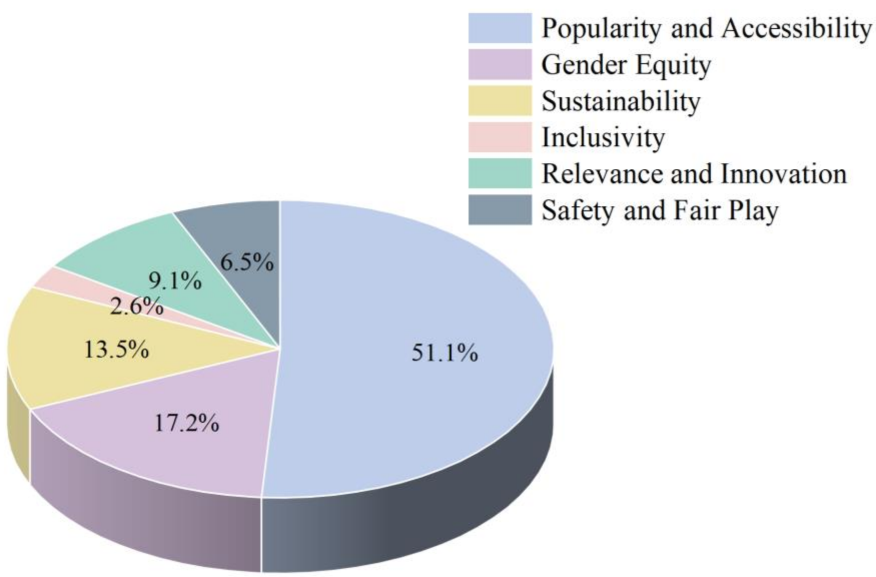

Figure 8. Weight of Each Criteria for Japan

5.2.2 TOPSIS II

As stated in Section 5.1.2, TOPSISI yields the final score of each criterion. Then, all 52 disciplines are used, including the 43 disciplines used to calculate weight and 9 that will be tested in Section 6. For each discipline, we calculate the distance to the positive and negative ideals, and the final score C_i , with i being the discipline. Since these weights and rankings are dependent on the hosting countries, results for 34 countries are calculated and presented in Appendix 8.

We present the disciplines that ranked 1st for 34 different countries in Figure 9. The countries that have the same discipline are painted the same color.

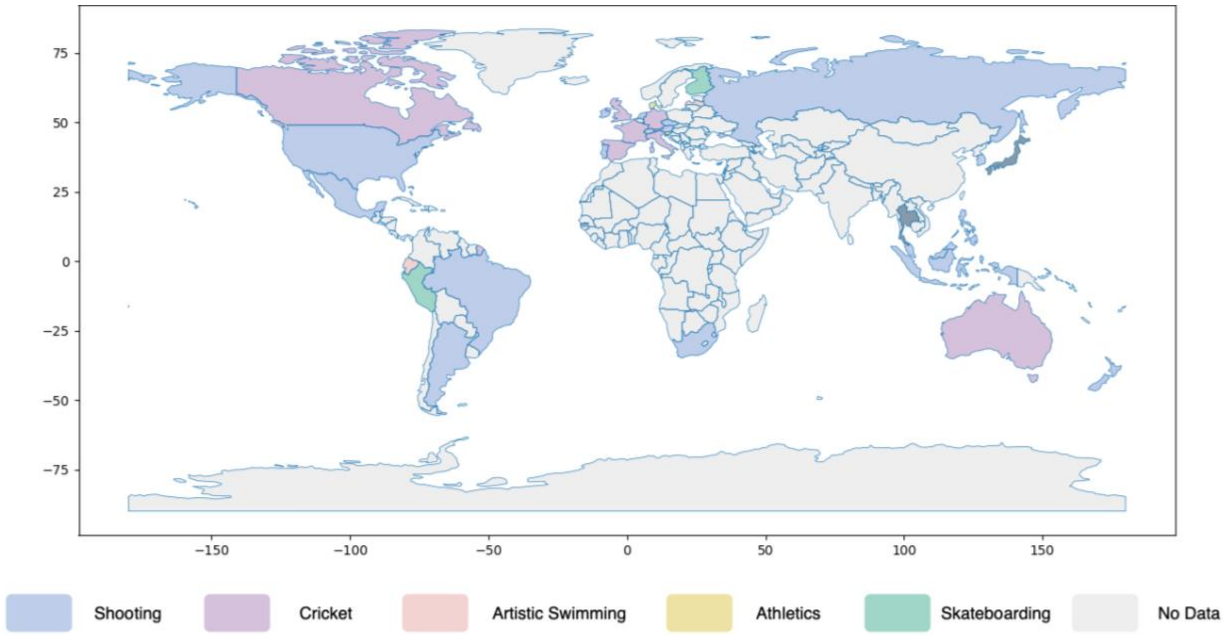

Figure 9. World Map of Disciplines Ranked First by Colors

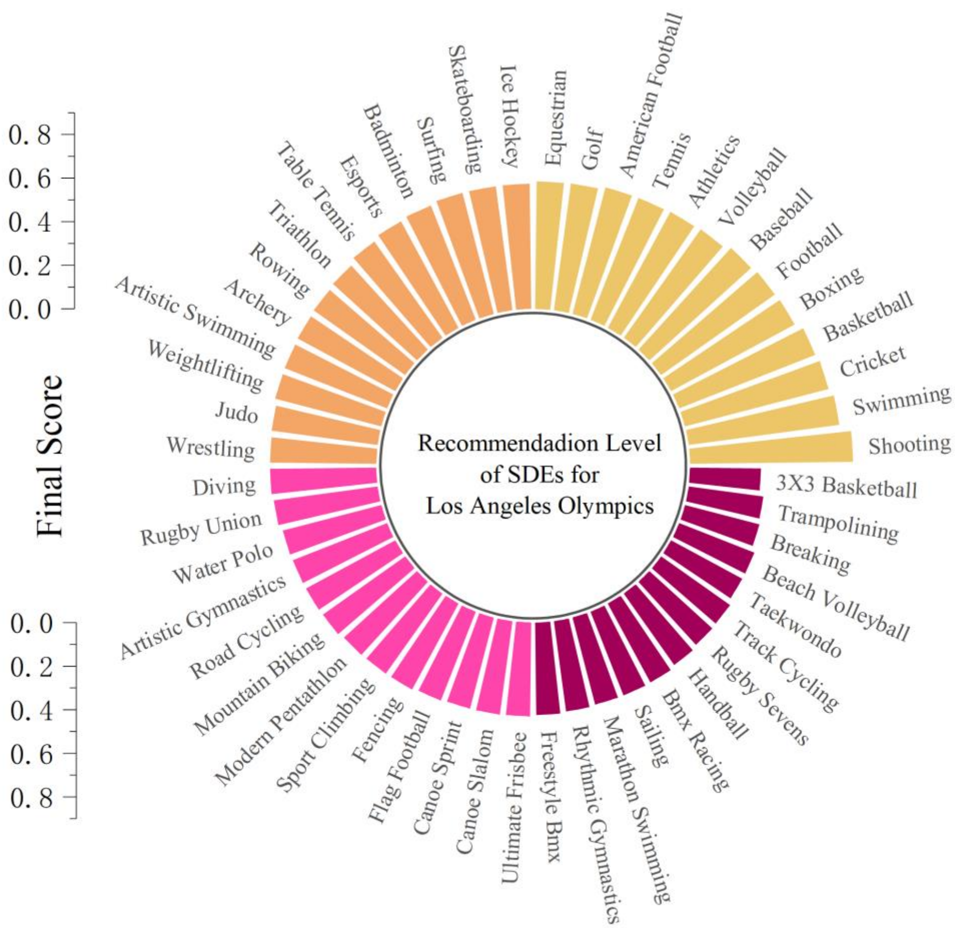

Figure 10. Final Evaluation Score for 52 Disciplines for 2028 Los Angeles Olympic

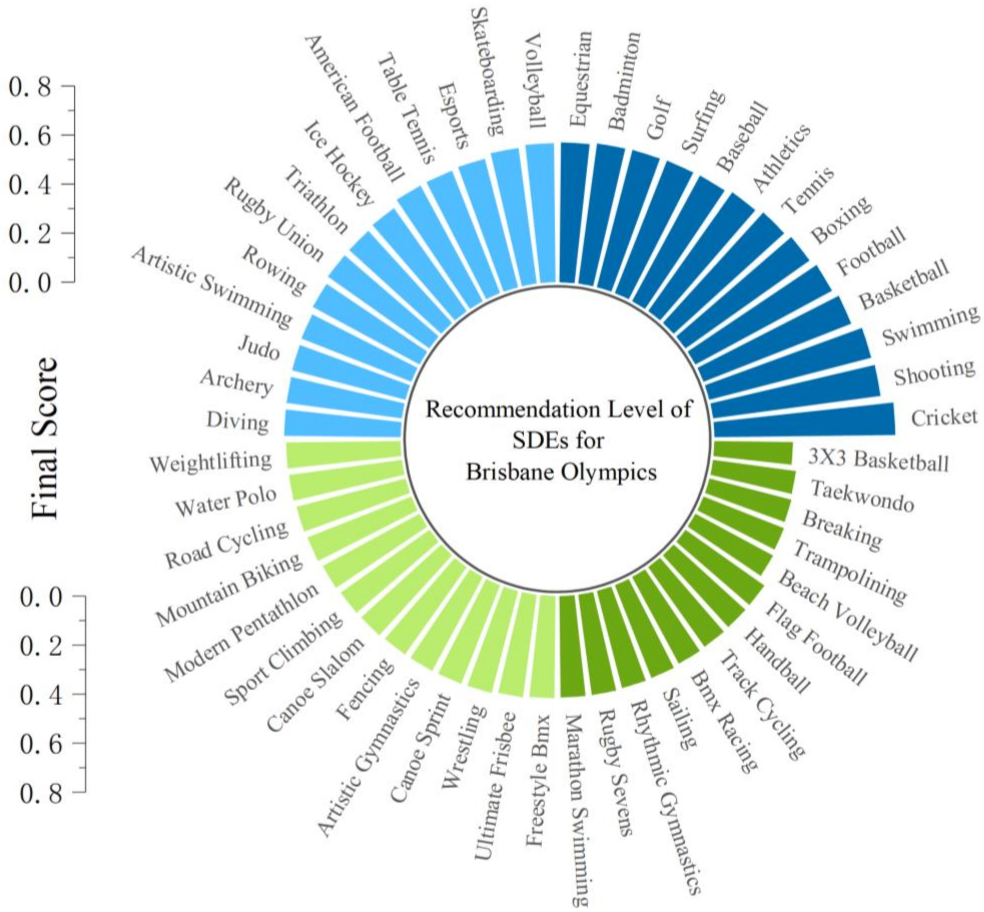

Figure 11. Final Evaluation Score for 52 Disciplines for 2032 Brisbane Olympic

The scores and recommendation level, with higher bars representing higher recommendation, for each discipline in United States are presented in Figure 10. The results for 2032 Brisbane Olympic in Australia are in Figure 11.

6 Model Implementation

The final output of this model is the ranking of all 43+9 untested and tested disciplines. The nine disciplines to be tested did not take part in determining the weights and building the model. They can be categorized to:

(1) Breaking, Baseball and Softball, and Flagged football: These are disciplines that are added or removed from the 2028 Los Angeles Olympic. The popularity data will be the popularity within the United States.

(2) Rowing, Fencing, and Artistic Gymnastic: These are disciplines that have continued to be in the Olympic since 1988.

(3) The final category is disciplines we suggest being introduced to the 2032 Brisbane Olympic, and the model is based on data in Australia.

For all three categories, the Criterion Popularity and Accessibility are based on data from the hosting country. Cultural and traditional preferences can lead to certain sports being more popular in the host country. For example, karate was introduced when Japan hosted the Olympics, and baseball and softball was introduced again when the United States hosted. This "home-field advantage" can influence the popularity of events and even affect the attention of a global audience.

The six disciplines from Category 1 and 2 shows varying performance on criteria, which allows diverse aspects of our model to be tested. For example, Baseball and Softball scored 0.91 on C5 (Relevance and Innovation) and 0.64 on C4 (Inclusivity), while Artistic Gymnastics scored 0.94 on C4 and Rowing with 0.66 on C5.

First, we collected data on six criteria and sub-criteria related to these nine disciplines. All data for these disciplines is collected and preprocessed in the same way as the data for the other 43 disciplines, as mentioned in Section 4. But we select specific country for Popularity data in Google Trends. Then, we input the data into the initial matrix $Formula$ (1). Since we already know the weights of each criterion and sub-criterion from Section 5, we can directly use TOPSIS to calculate the final scores.

These final scores will be compared against the ranking of other 43 untested disciplines. By observing their ranking, our recommendation level and decision of whether a discipline should be added are shown in Table 2. Disciplines that are ranked under the 25th percentile of the total 52 disciplines are rejected.

Ranking	**Recommendation Level**	**Decision**
1-13	Strongly Recommend	Accept
14-26	Recommend	Accept
27-39	Retain	Accept
40-52	Does not recommend	Reject

Table 2. Ranking and Corresponding Recommendation Level

Section 6.1 validates the decision of inclusion, removal and continuous presence of disciplines in the first two categories, while Section 6.2 uses our model to support the introduction of new disciplines.

6.1 Testing Against SDEs

For Category 1, our results for 2028 Los Angeles Olympic are shown in Table 3. Breaking was first introduced in the Paris Olympics but was removed from the Los Angeles Olympics. According to our calculations, breaking ranks 50th for the Los Angeles Olympics, with a recommendation of "Not Recommended," validating the IOC's decision to remove it. Baseball and softball were removed from the Paris Olympics but reintroduced for the Los Angeles Games. Our model rank baseball and softball 7th, with a "Strongly Recommended" status, again supporting the IOC's decision. Flag football was introduced for the first time in the Los Angeles Olympics. Our calculations place it at 36th, with a "Retain" recommendation. Looking at the scores across the six criteria, flag football scores relatively high in C1 (Popularity and Accessibility) compared to other events at the "Retain" level, but lower in C2 (Gender Equity) and C6 (Safety and Fair Play). This suggests that the IOC may believe that flag football has the potential to improve in terms of safety measures and gender balance in its future development.

Name of Discipline	TOPSIS Score	Ranking	Recommendation Level	IOC Decision
Breaking	0.3525838871949216	50	Does not recommend	Reject
Baseball and Softball	0.6308426360212754	7	Strongly Recommend	Accept
Flagged Football	0.43884290612129145	36	Retain	Accept

Table 3. Decisions for Category 1 Disciplines

For Category 2, the TOPSIS score, ranking (R) and recommendation level (RL) of these disciplines for 2016 to 2032 Olympics are shown in Table 4. Rowing consistently ranks around 20th across the four Olympic Games we analyzed, yielding a "Recommend" result. Fencing and Artistic Gymnastics both rank around 30th, with a "Retain" recommendation. The decision for all three sports is to retain them, matching the IOC's decisions. This shows that these three sports are less affected by changes in the host country and enjoy relatively stable popularity across different geographical locations. Their consistent "Retain" decision validates their continued presence in the Olympics.

Name of Discipline	2020		2024		2028		2032	
	R	RL	R	RL	R	RL	R	RL
Rowing	25	Recommend	24	Recommend	21	Recommend	22	Recommend
Fencing	36	Retain	36	Retain	35	Retain	32	Retain
Artistic Gymnastic	28	Retain	38	Retain	30	Retain	35	Retain

Table 4. Decisions for Category 2 Disciplines

6.2 Proposing New SDEs

Considering innovation and diversity, we identify three new disciplines. Then, we tested them with our model. The name, score and ranking of these disciplines are presented in Table 5.

Name of Discipline	TOPSIS Score	Ranking	Recommendation Level
Frisbee	0.42282492068390637	38	Retain
E-sports	0.5416331529120427	16	Recommend
Cricket	0.7376487049045042	1	Strongly Recommend

Table 5. Decisions for Category 3 Disciplines

For the Brisbane 2032 Olympics, we anticipate the potential addition of three events. According to our calculations, cricket ranks 1st with a "Strongly Recommended" status. E-sports ranks 16th, yielding "recommend." Frisbee ranks 36th with a "Retain" recommendation. Cricket should be considered a top priority for inclusion, followed by e-sports. While frisbee's current data does not show particularly high popularity, its visibility has been growing rapidly in recent years [17]. By 2032, its actual score may increase, making it a potential third candidate for consideration.

7 Modified Model

7.1 Sensitivity Analysis

7.1.1 Quantifying Change in Evaluation

We use the Root Mean Square Error (RMSE) to interpret the sensitivity. This method punishes large residues and gives an interpretable scale for the error between the values. The smaller the RMSE means the model's predictions are closer to the actual values. RMSE is calculated through:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

n is the number of elements, in this case $n = 43$ for disciplines. y_i is the evaluation score for the ith discipline based on the modified model. $\hat{y}_i$is the score based on our original model.

7.1.2 Changing Disciplines

In our model, the 43 disciplines that determine the weights are all long-term existing Olympic sports. These sports show what an Olympic sport should be like. Our first sensitivity analysis focuses on changing these sports.

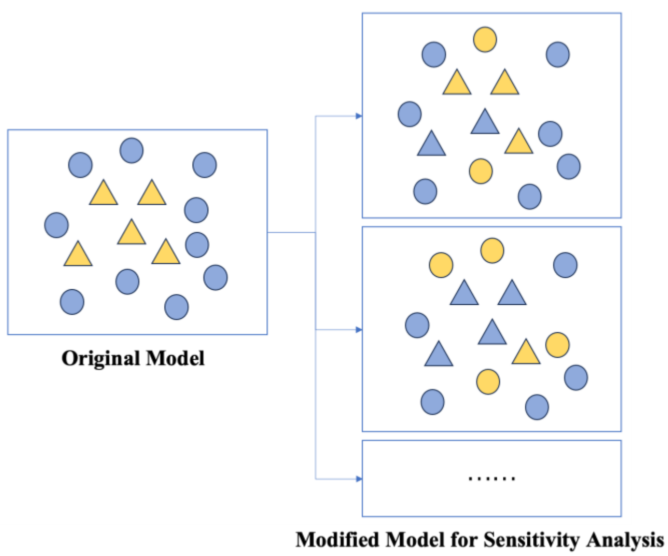

Figure 12. Abstraction of Randomized Change in Disciplines

Out of 52 total disciplines, randomly select 43 as the disciplines that determine the weights. By analyzing all the sports again, we calculate the RMSE of our original model and the modified version. Figure 12 is an abstraction of this step: the circles being 43 long-term existing Olympic disciplines; triangles being 9 disciplines used for evaluation; blue-colored items representing disciplines used to determine the weights; yellow-colored items as disciplines for verification.

Over 100 different sets of 43 disciplines, our model averaged an RMSE of 0.01031. Despite the disciplines being altered, the average change in the scores has barely varied. This shows that our model is robustness is high, and the fact that some disciplines have been altered does not affect the system. The code that provides this information is in Appendix 9.

7.1.3 Changing Inputs

By adding a random number between ± 0.05 to every single standardized value, the model's outputs only have a RMSE of 0.01801. This shows that our model provides a holistic view of the disciplines, and it cannot be altered by slight increases or decreases of the original data. Figure 13 is an abstract representation of this step, with the error bars showing ±0.05 and triangles symbolizing randomness. The code used is in Appendix 10.

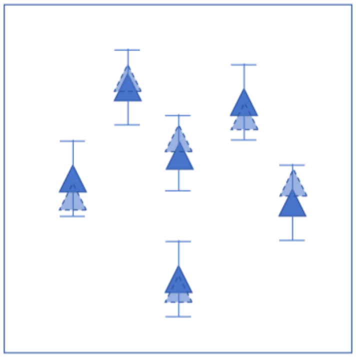

Figure 13: Abstraction of Randomized Change in Inputs

7.2 Extensions

In addition to modification made in Section 7.1, we may modify the model to make it dependent on time. In Assumptions 1 and 4, we assumed the criteria and data are irrelevant with changes due to time. We only addressed Popularity and Accessibility, by selecting data from the recent five years. In reality, data for all criteria— such as audience engagement, cost, and infrastructure availability—are dynamic and may vary significantly over time. These temporal changes affect the weights assigned to the criteria and, consequently, the final rankings. To take this into consideration, each factor can be modeled as a function of time using methods such as curve fitting or regression analysis. This approach would enable us to estimate the relationship between each criterion and time, then use these functions to predict future values for the factors. For example, we could analyze trends to forecast how disciplines might evolve in their fitness for the Olympics beyond 2032, therefore providing a more forward-looking evaluation.

Moreover, we only addressed disciplines and data in the Summer Olympics. The methods used for data collection and modeling may be used to evaluate disciplines in the Winter Olympics and the Paralympics. Additionally, we could modify the model's smallest evaluation unit to assess sports or events instead of disciplines. By integrating time-dependence and expanding the model's scope, we can enhance its versatility and applicability across different Olympic contexts.

8 Conclusion

In addition to being the first model that evaluate all six IOC criteria, our work has the following highlights:

(1) We collected extensive data covering 52 disciplines and 6 criteria, including both Olympic statistics and external data such as gender ratio statistics and global interest provided by Google Trends.
(2) We processed data through iteration, transformation and tailored normalization. we ensured that all factors were comparable and contributed equitably to the final evaluation score.
(3) Our model incorporated EWM with TOPSIS, which presents an objective ideal condition and compares each discipline with it, deducing an evaluation score. This allows a more holistic, objective and data-driven approach to decision-making.
(4) We generated 34 distinct sets of discipline ranking for different host countries, based on the evaluation score, which considers the impact of cultural differences on data.
(5) We validated our model through sensitivity analysis using Root Mean Square Error.

We tested our model on Breaking, Baseball and Softball, Flagged Football, Rowing, Fencing, and Artistic Gymnastics. The recommendations made by our model all matched the decisions of the IOC. We suggest, in order of recommendation, e-sports, cricket and frisbee, to be added to the 2032 Olympics in Brisbane.

To conclude, our work provides a comprehensive and robust model to evaluate discipline's fitness for the Olympic Games. Our model goes beyond the subjective metrics and incorporates diverse data sources. It offers a holistic, objective, and data-driven approach to evaluate disciplines. We welcome further explorations, such as evaluation dependent to time and broadening the scope to Winter and Paralympics. All in all, this model is a valuable tool for the IOC in shaping the future of the Olympic program, ensuring its continued relevance, global appeal, and alignment with the core values of the Olympic movement.

9 Letter

Maison Olympique,
1007 Lausanne,
Switzerland

Dear Members of the International Olympic Committee,

We are consulting team 14977, investigating Sports, Disciplines and Events (SDEs)'s fitness to the IOC's criteria. We aim to create an assessment model that evaluates all IOC criteria to provide an objective and data-driven approach in evaluating potential SDEs. We would like to present our findings by explaining: (1) our research, (2) the rationale of our model, and (3) our recommendation for SDEs addition.

We suggest specific factors for each IOC criteria. For example, we use Global Interest and Host Country's Regional Preference to identify the criterion Popularity and Accessibility. And we provide 20 yes or no questions for Relevance and Innovation. For all factors, we collect general public data that is changes for different hosting countries, varying from gender ratio statistics from previous studies to popularity provided by Google Trends, for each discipline — the smallest evaluation unit we chose.

Subsequently, we create a model that evaluates how well a discipline fit in the Olympic, by considering the weights, or the ability a factor can influence the decision, of each criterion. We find the positive ideal condition (what is desired) of a discipline, which is when every criterion has the greatest score, and the negative ideal condition (what we do not want). Each discipline's score is then compared to see how they are to what we want.

Our model produces 34 sets of scores ranking, based on 34 different countries, for all disciplines currently in the Olympic. Our model evaluates which discipline should be added to the Olympic Games. Also, it suggests disciplines that may be removed from the Olympic. We decided that if a new discipline is ranked above 75th percentile, for the model of the hosting country, then it strongly recommended to be added. If the discipline is ranked below the 25th percentile is rejected.

Based on this standard, we tested our model on Breaking (will be removed in 2028), Flagged Football (will debut in 2028), Baseball and Softball (will return in 2028); and Rowing, Fencing, and Artistic Gymnastic that have continuously been in the Olympic. Our results are consistent with the decisions of the committee.

We would also like to suggest three SDEs that have the potential to be added to the 2032 Olympics in Brisbane. They are, in order of recommendation:

- Cricket: having ranked 1st, cricket is strongly recommended.
- E-sports: having ranked 16, e-sports is recommended.
- Frisbee: having ranked 38, frisbee is marked "retain," but with popularity growing rapidly in recent years, is still recommended.
-

Hope our findings are useful to promote global and fair Olympic Games!

Sincerely,
HiMCM Olympic Consultant Team 14977

References

[1] Thomson, A., Kennelly, M., & Toohey, K. (2020). A systematic quantitative literature review of empirical research on large-scale sport events' social legacies. Leisure Studies, 39(6), 859-876.

[2] Müller, M., Wolfe, S. D., Gaffney, C., Gogishvili, D., Hug, M., & Leick, A. (2021). An evaluation of the sustainability of the Olympic Games. Nature sustainability, 4(4), 340-348.

[3] Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the sustainable development goals. Nature sustainability, 2(9), 805-814.

[4] Celik, E., Gul, M., Aydin, N., Gumus, A. T., & Guneri, A. F. (2015). A comprehensive review of multi criteria decision making approaches based on interval type-2 fuzzy sets. Knowledge-Based Systems, 85, 329-341.

[5] Byers, T., Hayday, E. J., Mason, F., Lunga, P., & Headley, D. (2021). Innovation for positive sustainable legacy from mega sports events: virtual reality as a tool for social inclusion legacy for Paris 2024 paralympic games. Frontiers in sports and active living, 3, 625677.

[6] Loland, S. (2006). Olympic sport and the ideal of sustainable development. Journal of the Philosophy of Sport, 33(2), 144-156.

[7] Wu, J., & Zhang, Y. (2008). Olympic Games promote the reduction in emissions of greenhouse gases in Beijing. Energy Policy, 36(9), 3422-3426.

[8] Yang, J. J., Lo, H. W., Chao, C. S., Shen, C. C., & Yang, C. C. (2020). Establishing a sustainable sports tourism evaluation framework with a hybrid multi-criteria decision-making model to explore potential sports tourism attractions in Taiwan. Sustainability, 12(4), 1673.

[9] Haake, S. J. (2009). The impact of technology on sporting performance in Olympic sports. Journal of Sports Sciences, 27(13), 1421-1431.

[10] Google Trends. "Title of the Specific Page or Content If Applicable." Accessed November 12, 2024. https://trends.google.com.

[11] Xu, Q., Billings, A. C., Scott, O. K., Lewis, M., & Sharpe, S. (2019). Gender differences through the lens of Rio: Australian Olympic primetime coverage of the 2016 Rio Summer Olympic Games. International Review for the Sociology of Sport, 54(5), 517-535.

[12] Pons-Villanueva, J., Seguí-Gómez, M., & Martínez-González, M. A. (2010). Risk of injury according to participation in specific physical activities: a 6-year follow-up of 14 356 participants of the SUN cohort. International journal of epidemiology, 39(2), 580-587.

[13] Kim, W., Jun, H. M., Walker, M., & Drane, D. (2015). Evaluating the perceived social impacts of hosting large-scale sport tourism events: Scale development and validation. Tourism management, 48, 21-32.

[14] Hellal, M. (2023). Technological Fair Play: An Ethical Framework for Olympic Sports.

[15] Aydoğdu, E., Güner, E., Aldemir, B., & Aygün, H. (2023). Complex spherical fuzzy TOPSIS based on entropy. Expert Systems with Applications, 215, 119331.

[16] Hwang, C.L.; Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. New York: Springer-Verlag.

[17] "Participants in ultimate frisbee in the U.S. from 2006 to 2017." Statista, July 2018, https://www.statista.com/statistics/191967/participants-in-ultimate-frisbee-in-the-us-since-2006/. Accessed November 12, 2024.

Appendix

[1] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/3%20Sustainability.json
[2] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/5-1%20Relevance.json
[3] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/5-2%20Innovation.json
[4] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/intermediate_data/popularity.json
[5] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/1-2%20Popularity.json
[6] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/1-1%20Local.json
[7] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/input_data/246.csv
[8] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/code4model/reg.py
[9] https://github.com/SinoHero/Olympics-sports-assessment/tree/main/implementation_on_various_countries
[10] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/code4model/sensitivity.py
[11] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/code4model/sensitivity3.py
[12] https://github.com/SinoHero/Olympics-sports-assessment/blob/main/weights.txt