Problem Chosen

A

2024

HiMCM

Summary Sheet

The essence of Olympism is to encourage interactions within the human society, stand on guard on human dignity, and advocate for peace through sports. However, there are countless **Sports, Disciplines, and Events (SDEs)** around the world, each equipped with its own strengths to accomplish this goal. To determine which SDEs are the best fit to be added to both the 2032 Brisbane Olympics, and to explain the longevity or the absence of certain SDEs in the Olympic Games, we developed a mathematical model that ranks each SDE to help us evaluate which SDEs fit Olympism the most.

Using the **Analytical Hierarchy Process (AHP)** and **Shannon Entropy Method**, we computed a score and rank for all current Olympic SDEs, as well as judged whether new SDEs should be added to the Olympics. After acquiring our results, we then drafted a letter to the IOC, presenting our findings, as well as our recommendations for SDEs that we found suitable to be added: including Karate, Pickleball, and Ultimate Frisbee™.

To rank the SDEs, we used seven indices in our model: **Geopolitical Inclusivity**, **Gender Equity**, **Safety**, **Sustainability**, **Popularity**, **Relevance and Innovation**, and **Fair Play**. After we had developed our model, we found data for each of our indices. We calculated the Inclusivity Index based on the number of National teams participating. Next, the Gender Equity Index was measured by the proportion of male and female athletes. Then, the Safety Index was determined by the rates of injury. After that, the Sustainability Index was calculated based on carbon emissions related to the facility where the SDE is held. We used an API program to pull data from Google Trends™, determining the popularity of keywords related to the SDE for our Popularity Index. Afterwards, a ranking system was used to decide the score of the Relevance and Innovation index, with a score being assigned to each SDE from $^0/_5$ to $^5/_5$ based on the extent of its innovation(s). Finally, the Fair Play Index was measured by the rates of doping among athletes in said SDE.

After collecting and processing the necessary data, we normalized the values of each index and applied weights obtained from both the AHP and the Shannon Entropy Model. We then summed up the weighted scores of the indices to get a final score for each SDE. Ultimately, the SDE's are ranked by score, from best to worst fit for the Olympics, and **Sensitivity Analysis** was performed on the model based on our data.

The most creative aspect of our model for ranking each SDE's compatibility with the Olympics was the integration of AHP, which is subjective, and the Shannon Entropy method, which is objective but sometimes not precise. By combining these two methods, we can balance out the inherent bias of both models and develop the most logical and optimal weighting for each index.

Keywords: AHP, Shannon Entropy Method, Sensitivity Analysis

Contents

List of Figures

List of Algorithms

1 Introduction

1.1 International Olympic Committee

Founded at the first Olympic Congress in Paris in 1894, the **International Olympic Committee (IOC)** is the leader of the Olympic Movement and the guardian of the **Olympic Games**.[1] Under the banner of the IOC, there has been 30 Olympic Games as of 2024. **Brisbane 2032**, officially known as *Games of the XXXV Olympiad*, will be the 32^{nd} Olympic Games ever hosted.[1]

The purpose of the Olympic Games is to provide the means of peaceful interactions between nations, to safeguard human dignity, and to promote equal rights no matter the background. There are countless sports around the world, each with its own uniqueness and strength. However, the resources of the IOC and the host countries are limited. Therefore, it is important to decide which **sports, disciplines, and events (SDEs)** are to be included.

To assist the IOC in evaluating the SDEs, a mathematical model has been developed to evaluate sports based on core Olympic values. The model will test the SDEs against the **Criteria for Sports Inclusion** to provide reasonable and quantified recommendations.

1.2 Problem Restatement

1. Decision making is challenging when multiple factors have to be considered. One reason is that cognitive error is always present. In order to reduce the amount of cognitive error involved, two mathematical sub-models were jointed, in which one is more subjective and another is purely objective.

2. The aforementioned model is then utilized to evaluate the SDEs and rank them based on scores. An Olympic qualification bound was then determined. Then, the scores of the SDEs recently removed or added to the Olympics will be computed. The this model's output will then be compared to the actual decision made by the IOC.

3. Using these results, a letter will be written to the IOC presenting the model. This letter will be written for persons with no prior knowledge about modeling, such that the persons involved may make good use of this paper.

1.3 Assumptions and Justifications

1. **Assumption** : The number of continents with at least one member country in the specific sports governing body recognized by the IOC need not matter in regards to the diversity, equity and inclusivity of the geopolitical reality of such sport, as long as such number exceeds or is equal to the number of four.
   Justification : The number of continents has a narrow range of $4 \sim 5$. Compared to the number of countries that has a range of $75 \sim 227$, the number of continents provides minimal statistical significance, thus an inadequate parameter for analysis.

[1] three Olympic Games were canceled due to the world wars

2. **Assumption** : Each sport is responsible for all the economical costs and environmental impacts in all venues where such sport is to be hosted, even if the aforementioned venues are to be shared with other sports events.
   Justification : For our future generations, reuse of venues is strongly encouraged. However, this leads to a very philosophical question about causality. If sport A and sport B use the same venues with a cost of C, does A *cause* C? For the purposes of this model, sport A is totally responsible for the Cost of the venue, as A is a sufficient condition of C.

3. **Assumption** : Olympic standard venues of the same type, provided that they have been recently constructed,[2] have negligible difference in the inflation-adjusted cost, environmental impact, and other parameters.
   Justification : Olympic venues can be designed in different styles, often utilizing different materials with different costs and environmental footprints, but analyzing the architectural values and costs is beyond the nature of sports itself. To provide a fair analysis of all sports alike, architectural consideration will not be evaluated.

4. **Assumption** All Olympic standard venues constructed recently are in accordance with the regulations and standards of sports hosted, and all outdoor sports are hosted in similar conditions, implying a similar injury rate.
   Justification During *Paris 2024*, concerns about the water quality in which athletes competed in did exist.[2] However, analyzing the safety of the venues is beyond the scope of the topic, which is to evaluate the sports themselves. Therefore, it is assumed that all venues are within safety standards, and as such the injury rate of a given sport is only related to the nature of such sport.

5. **Assumption** All injuries, as long as it leads to one or more day off, are considered severe.
   Justification Since all injuries can and do lead to a possibility of an athlete's reduction of performance and basic function, sometimes permanently, it is important to consider all injuries when evaluating safety.

6. **Assumption** Carbon emission of constructing an arena is dependent on the amount of concrete.
   Justification The majority of carbon emission comes from the manufacturing of concrete. The inherit chemistry of calcium carbonate, the core ingredient of concrete, is responsible for 60% of the carbon emission,[3] which is not improvable by any means. Therefore, it can be safely assumed that the amount of carbon emission is mainly dependent of the amount of concrete used in construction.

7. **Assumption** The amount of concrete used in the construction of an arena is dependent on its type and its area.
   Justification As in assumption 3, the purpose of the model is not to judge the architectural design of the arena. Considering that there are meaningful differences in

[2] within twenty years, inclusive

the nature of arenas for different sports, only the arena's type and capacity will be considered.

8. **Assumption** Renovation for the purpose of upgrading to Olympic standard does not have a significant carbon footprint.
   Justification While it is true that renovation would inevitably lead to some carbon emission, it is minuscule if compared to constructing a new one. Thus it would not be taken into account.

9. **Assumption** All sports venues will be used and maintained everyday in the time span of the Olympic.
   Justification It is expected that all resources will be used, especially considering there are few venues qualified for the Olympic Games.

10. **Assumption** Google Trends™ can accurately represents global searches.
    Justification Since Google is the most used search engine by a large margin[4], and that search engine is generally not correlated to a preference in sports, it can be safely assumed that Google Trends depicts a certain sport's popularity accurately.

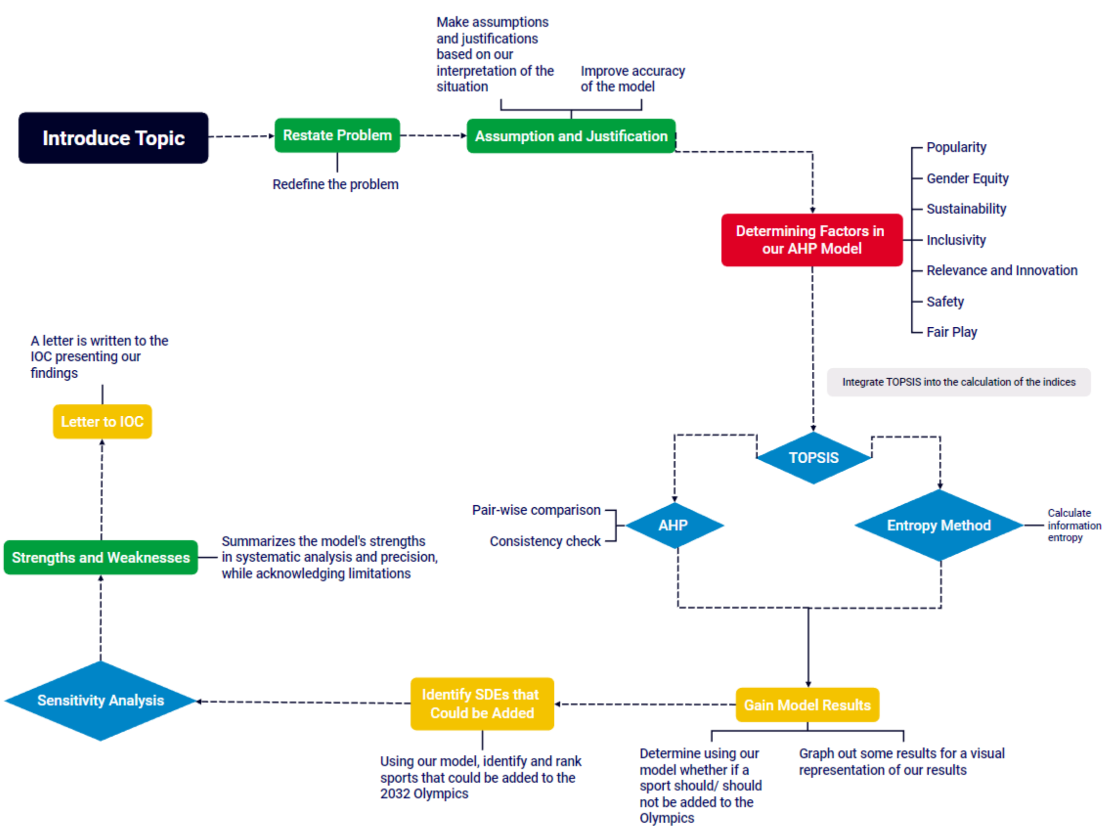


Figure 1: Work Structure

2 Table of Variables

Variables	Descriptions (Unit)	[1]	[2]	[3]
I_{GInc}	Geographical Inclusivity Index			D
v_{State}	Number of Eligible Countries in a Sport	□	V	D
v_{Ctn}	# of Continents w. min. 1 Eligible Country	□	V	D
I_{Gen}	Proportional Gender Representation Index			D
Δv_{Gen}	Diff. between Proportion of Genders	□	V	D
v_{GenM}	Proportional Representation of Male Athletes	□	V	D
v_{GenF}	Proportional Representation of Female Athletes	□	V	D
I_{Safe}	Safety Index			P
v_{Inj}	Injury Rate	□	V	P
I_{Sus}	Sustainability Index			D
v_{Area}	Area of a Sport Venue (m^2)	□	V	D
v_{Cnst}	Emission of Construction (t/m^2)	□	V	D
v_{Mn}	Emission of Maintenance ($t/m^2 \cdot \text{Day}$)	□	V	D
v_{Grn}	Positive Environmental Impact ($t/m^2 \cdot \text{Day}$)	□	V	D
v_{Day}	Duration of the Olympic Games (Day)	□	C	D
v_{NetEm}	Net Emission of a Sport Venue (t)	□	V	D
I_{AuPop}	Popularity Index - Australian Model			P
I_{GlbPop}	Popularity Index - General Model			P
v_{GlbAvg}	Average Popularity - Overall	□	V	P
v_{OlAvg}	Average Popularity - Olympic	□	V	P
v_{AuAvg}	Average Popularity - Australian	□	V	P
i_{GlbPop}	Overall Popularity Subindex			P
i_{OlPop}	Olympic Popularity Subindex			P
i_{AuPop}	Australian Popularity Subindex			P
I_{Inn}	Innovation Index	■	V	D
I_{Fair}	Fair Play Index			D
v_{dope}	Proportion of Athletes Doping	□	V	D
w_{AHP}	Weight of AHP Model			
w_{Entropy}	Weight of Shannon Entropy Model			
w_{final}	Weight of the Final Model			
s_{GlbM}	Score of the General Model			
s_{AuM}	Score of the Australian Model			

Table 1: Table of Variables Used and their Descriptions

[1] □: Quantity, ■: Quality
[2] C: Constant, V: Variable
[3] D: Deterministic, P: Probabilistic

The type of variables and indices left blank are to be understood to be calculated, and therefore, varying quantities.

3 Index Extraction and other Calculations

3.1 Geographical Inclusivity Index

"The five-ring logo [symbolizes] the five continents of Africa, the Americas, Asia, Europe and Oceania united by Olympism and in which - together with the white flag fabric - the colors of the flags of all then-known nations are to be found."[5] Under one banner, nations from across the globe are represented, allowing individual athletes to compete together no matter the background. As such, geographical inclusivity is certainly important to the spirit of the Olympic Games.

The geographical inclusivity index I_{GInc} is a quantitative measurement of the diversity and inclusivity of the sports evaluated. This index is computed by the number of member countries[3] v_{State} of the IOC-recognized sports governing body, and the number of continents v_{Ctn} of which at least one of the aforementioned countries are located.[4]

Algorithm 1 Pseudocode for Geographical Inclusivity Index

if $(v_{\text{Ctn}} < 4)$ **or** $(v_{\text{State}} < 75)$ **then**
  $I_{\text{GInc}} \leftarrow 0$
else
  $I_{\text{GInc}} \leftarrow (v_{\text{State}} - 74) \div (\max v_{\text{State}} - 74)$
end if

If the geographical inclusivity requirements, *i.e.* a minimum v_{State} of 75 and a minimum v_{Ctn} of 4, are not achieved, this algorithm returns 0 to I_{GInc}. If such requirements are achieved, a unit-interval value will be returned to I_{GInc} based on the min-max linear normalization dependents of v_{State} and $\max v_{\text{State}}$.[5]

3.2 Proportional Gender Representation Index

"Gender equality, inclusion and diversity are Fundamental Principles of Olympism in the Olympic Charter and central to fulfilling the IOC's and the Olympic Movement's vision of building a better world through sport."[7] In one of the largest event of the world, representation of both genders is undoubtedly essential. It is therefore an important criterion to consider as well.

The proportional gender representation index I_{Gen} is a quantitative measurement of the equity of representation among both genders. This index is computed by the proportion of male athletes v_{GenM} and female athletes v_{GenF}.

Algorithm 2 Pseudocode for Proportional Gender Representation Index

$\Delta v_{\text{Gen}} \leftarrow |v_{\text{GenM}} - v_{\text{GenF}}|$ ▷ Difference between Proportion of Genders
$I_{\text{Gen}} \leftarrow 1 - \Delta v_{\text{Gen}} \div (\max \Delta v_{\text{Gen}})$

[3] includes non-sovereign territories and states with limited recognition
[4] includes Africa, Americas, Asia, Europe, and Oceania
[5] As of 2024-11-08, the sports governing body with the most members is ITTF, at a number of 227.[6]

The difference between the proportion of genders is stored as Δv_{Gen}. Unlike the case for I_{GInc} and v_{State}, in which a higher v_{State} number corresponds to a better I_{GInc} score, a negative relation between the score I_{Gen} and the Δv_{Gen} should be present. Therefore, a unit-interval value will be returned based on the linear normalization followed by a positive transformation, based on Δv_{Gen} and $\max \Delta v_{\text{Gen}}$.

3.3 Safety Index

Despite advancing technologies and medical knowledge, injuries are still prevalent, sometimes accidental, sometimes intentional; sometimes inevitable, sometimes preventable. To encourage sports governing committees to implement safety measures, safety must be considered in this model.

The safety index I_{GInc} is a quantitative measurement of the safety of both players and staff involved. This index is solely based on the proportion of injuries[6] v_{Inj} per person per day.

Algorithm 3 Pseudocode for Safety Index

$I_{\text{Safe}} \leftarrow 1 - v_{\text{Inj}} \div \max v_{\text{Inj}}$

In the above algorithm, v_{Inj} undergoes linear normalization then a positive transformation, and the result will be returned to the safety index I_{GInc}.

3.4 Sustainability Index

In cities where the Olympics are to be held for the first time, pre-existing infrastructure is often unqualified to host the largest sport event in the world.[7] As a result, new infrastructure will often have to be built to accommodate Olympic standard fields, pools, and auditoriums. These factors lead to the Olympics being a substantial environmental burden. In this subsection, the carbon emissions of construction, maintenance, and positive impacts of said infrastructure will be evaluated.

3.4.1 Carbon Emission of Constructing Olympic Infrastructure

As per assumption 6 and 7, carbon emission is dependent of the type and area of the sports venue only. Although there are plans[8] already, this model will utilize an estimated area to exanimate possible sports not currently included in the Queensland[8] urban planning of the 2032 Olympics. The total amount of carbon dioxide emitted through construction is calculated by multiplying its area v_{Area} by the amount of carbon dioxide it emits per meter squared v_{Cnst}.

[6] that leads to at least one day off

[7] by investment

[8] Brisbane is located in the state of Queensland

3.4.2 Carbon Emission of Maintenance

The environmental burden of maintenance during a short time frame[9] is relatively minor when compared to that of the construction process, but still a substantial amount that cannot be ignored. The total amount of carbon dioxide emitted when a sports venue undergoes maintenance is calculated by it area v_{Area} multiplied by the emission of maintenance per area per time v_{Mn} multiplied by $v_{\text{Day}} = 16$ Days of maintenance, as per assumption 9.

3.4.3 Positive Environmental Impact

While building new arenas and infrastructure may appear to have only negative environmental impacts, it is important to consider some positive impacts that certain venues of sports could offer.[10] The total Positive environmental impact of a sports venue is calculated by its area v_{Area} multiplied by the amount of carbon dioxide absorbed/neutralized per area per time v_{Grn} multiplied by $v_{\text{Day}} = 16$ Days.

3.4.4 Final Result of Sustainability Index

The sustainability index I_{Sus} is determined by the area v_{Area}, the estimated carbon emission per area of a venue v_{Cnst}, the carbon emission of maintenance per area per time of a venue v_{Mn}, and the positive environmental impacts per area per time of a venue v_{Grn}, which can vary depending on the type of infrastructure used by sport.

Algorithm 4 Pseudocode for Sustainability Index

$v_{\text{NetEm}} \leftarrow 0$
$v_{\text{Day}} \leftarrow 16$ ▷ Duration of the Olympic Games
$v_{\text{NetEm}} \leftarrow v_{\text{NetEm}} + v_{\text{Area}} \times v_{\text{Cnst}}$ ▷ Environmental Impact of Construction
$v_{\text{NetEm}} \leftarrow v_{\text{NetEm}} + v_{\text{Area}} \times v_{\text{Day}} \times v_{\text{Mn}}$ ▷ Environmental Impact of Maintenance
$v_{\text{NetEm}} \leftarrow v_{\text{NetEm}} - v_{\text{Area}} \times v_{\text{Day}} \times v_{\text{Grn}}$ ▷ Positive Environmental Impact
$I_{\text{Sus}} \leftarrow 1 - (v_{\text{NetEm}} - \min v_{\text{NetEm}}) \div (\max v_{\text{NetEm}} - \min v_{\text{NetEm}})$

The net emission v_{NetEm} is calculated by the emission of construction $v_{\text{Area}} \times v_{\text{Cnst}}$, plus the emission of maintenance $v_{\text{Area}} \times v_{\text{Day}} \times v_{\text{Mn}}$, subtracted by the positive environmental impact $v_{\text{Area}} \times v_{\text{Day}} \times v_{\text{Grn}}$. After the calculation, v_{NetEm} undergoes min-max linear normalization and and a positive transformation, which will become the sustainability index I_{Sus}.

3.5 Popularity Index

Popularity is undoubtedly the most important aspect in this model. After all, for the message of Olympism to spread, popular sports are required. Google Trends™ is a useful tool to evaluate the popularity of a sport, given the popularity of Google as a search engine (see figure 2). Data could be easily gathered by an API[11] program, which can be seen at section

[9] that of the duration of the Olympic Games
[10] An environmentally friendly golf course, for example.
[11] Application Programming Interface

10. Three types of popularity data were used as sub-factors for the popularity index: Overall popularity, Olympic popularity, and Australian Popularity.

3.5.1 Overall Popularity Subindex

The overall popularity factor i_{GlbPop} is computed from the search rate of a sport on Google over the past 20 years in proportion to a constant maximum point. The arithmetic average of this proportion is v_{GlbAvg}.

Algorithm 5 Pseudocode for Overall Popularity Subindex

$$i_{\text{GlbPop}} \leftarrow (v_{\text{GlbAvg}} - \min v_{\text{GlbAvg}}) \div (\max v_{\text{GlbAvg}} - \min v_{\text{GlbAvg}})$$

The overall popularity subindex i_{GlbPop} is computed by the min-max linear normalization of the proportion v_{GlbAvg}.

3.5.2 Olympic Popularity Subindex

The Olympic popularity subindex i_{OlPop} is computed from the arithmetic average v_{OlAvg} of the search rate of a sport *during* the Olympic Games in proportion to the constant maximum point. This subindex takes into account that some sports may only be popular during the Olympics. The calculation is similar to the subindex before.

Algorithm 6 Pseudocode for Olympic Popularity Subindex

$$i_{\text{OlPop}} \leftarrow (v_{\text{OlAvg}} - \min v_{\text{OlAvg}}) \div (\max v_{\text{OlAvg}} - \min v_{\text{OlAvg}})$$

3.5.3 Australian Popularity Subindex

Similar to the overall popularity subindex, the Australian popularity subindex also uses the arithmetic average of the search rate in a duration of 20 years. However, the data is collected specifically in Australia. When calculating the popularity of a sport with a specific geographical location of the Olympics, this subindex can be utilized. In Olympic Games past *Brisbane 2032*, this model can be reused by changing the country for this subindex. The calculation is similar to the previous two.

Algorithm 7 Pseudocode for Australian Popularity Subindex

$$i_{\text{AuPop}} \leftarrow (v_{\text{AuAvg}} - \min v_{\text{AuAvg}}) \div (\max v_{\text{AuAvg}} - \min v_{\text{AuAvg}})$$

3.5.4 Popularity Index (Australian)

The Australian model is used for calculating the popularity of sports for *Brisbane 2032* since it is location-specific. This model can also be modified to apply to future countries hosting the Olympics by changing the Australian popularity subindex to that of another countries'.

Algorithm 8 Pseudocode for Australian Popularity Index

$I_{\text{AuPop}} \leftarrow i_{\text{GlbPop}} \times 0.528 + i_{\text{OlPop}} \times 0.333 + i_{\text{AuPop}} \times 0.140$

To be more rigorous with the above calculation, a small scale AHP is used to determine the weighting of the three subindex involved. The table is as follows:

①	Overall Popularity
②	Olympic Popularity
③	Australian Popularity

Table 2: Factors Considered

	①	②	③
①	1	2	3
②	$1/2$	1	3
③	$1/3$	$1/3$	1

Table 3: Filled AHP table

For more information regarding AHP, see section 4.1.

3.5.5 Popularity Index (General)

Different from the Australian Popularity Index, the general model can compute the popularity of a sport globally. This index may be used when it is unclear where the future Olympic will be hosted, and there is still a need to determine which SDEs to be added or removed. Since the General model involved only two subindices, mathematical modeling is not needed. Instead, the weight of the overall popularity subindex and Olympic popularity subindex are set to 60% and 40%, respectively.

Algorithm 9 Pseudocode for General Popularity Index

$I_{\text{GlbPop}} \leftarrow i_{\text{GlbPop}} \times 3/5 + i_{\text{OlPop}} \times 2/5$

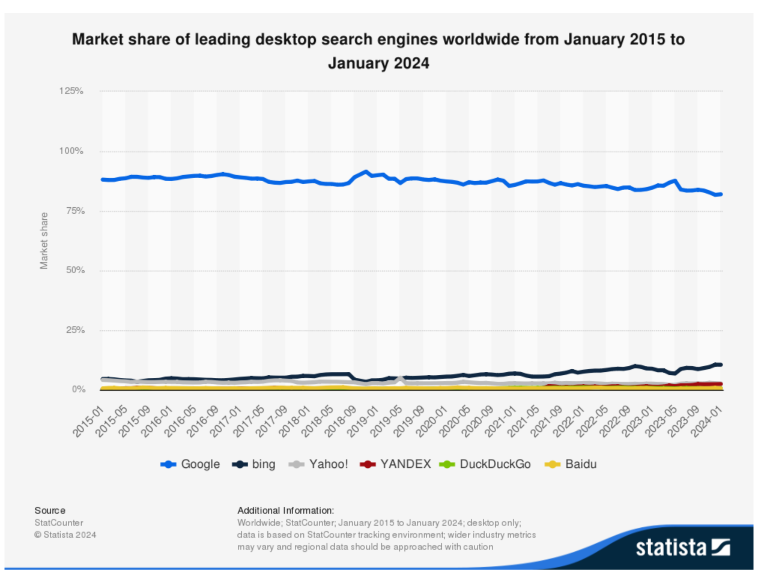


Figure 2: Global Desktop Search Engines Market Share[4]

3.6 Relevance and Innovation Index

Due to the qualitative nature of Innovation, as well as the lack of quantifiable metrics, a judgment metric is used. The innovation Index I_{Inn} is judged by the metric below:

I_{Inn}	Standard
$^0/_5$	• Complete absence of innovation
$^1/_5$	• Innovations lack any noticeable impact • Innovations does not solve current problems • Innovation does not relate to virtual sports
$^2/_5$	• Innovations have small but observable impacts • Innovations do not solve current problems • Virtual sports are applied at the beginning level, with little impact
$^3/_5$	• Innovations bring noticeable impacts • Innovation dedicated to solve current problems • Innovation involves virtual sports partially and has noticeable impact
$^4/_5$	• Innovations have outstanding impacts • Innovations have solved current problems • Innovation connects with virtual sports and has an outstanding impact
$^5/_5$	• Innovations have powerful impacts on the status quo • Innovation overcomes major barriers • Innovation may incorporate virtual sports and has a significant impact

Table 4: Metric for Innovation Index

3.7 Fair Play Index

The Fair Play index I_{Fair} focuses on measuring the rate of doping incidents in athletes, which save for being highly unfair to the competition, is also the main concern for the IOC when it comes to fair play. The Fair Play Index I_{Fair} is computed from fractions v_{dope} of the number athletes found doping over the total number of athletes of the SDE computed.

Algorithm 10 Pseudocode for Fair Play Index

$I_{\text{Fair}} \leftarrow 1 - v_{\text{dope}} \div \max v_{\text{dope}}$

4 Evaluation Model

4.1 The Analytical Hierarchy Process (AHP)

$$[a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{17} \\ a_{11} & a_{22} & \cdots & a_{27} \\ \vdots & \vdots & \ddots & \vdots \\ a_{71} & a_{72} & \cdots & a_{77} \end{bmatrix}$$

Figure 3: Notation of Matrices

The Analytical Hierarchy Process (AHP) is based on the premise that intransitivity is not only supposed to be allowed but even expected in the mathematical modeling of human evaluation and decision prediction. This is because humans are inherently irrational and 'predictably irrational'. Thomas L. Saaty, the inventor of this process claimed that humans compare criteria pairwise, and then aggregate the results. Importantly, "pairwise comparisons are performed at the neural level, the division algebra of octonions, in which commutativity and associativity are not satisfied."[9] In this particular model, given the relatively minuscule amount of indices, utilizing a simple model would be ideal. Therefore, it is decided that AHP will serve as the principal modeling method.

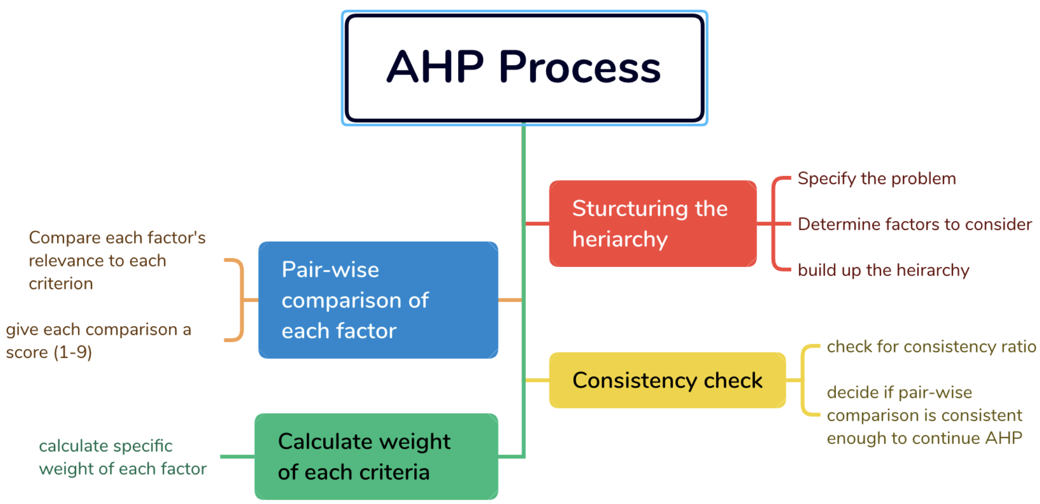


Figure 4: This AHP Process

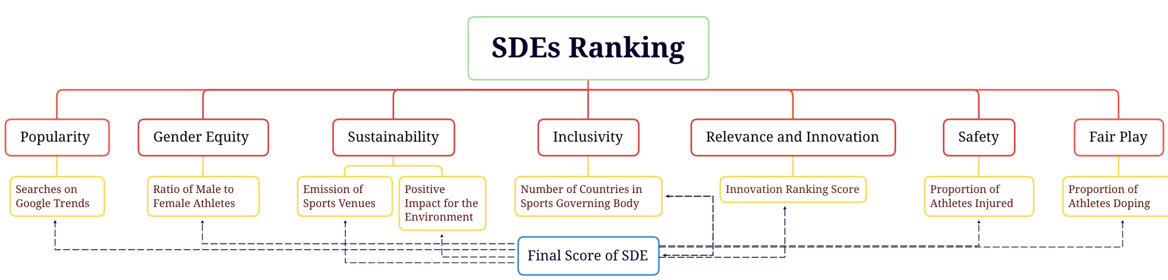


Figure 5: The AHP Structure of This Model

The method itself is not complicated. An individual element $a_{i\,j}$ of $[a_{i\,j}]$ is filled with a score as according to table 5. For example, if aspect ① is thought to be of higher importance than aspect ②, $a_{1\,2}$ would be set to 3.

Score	Meaning
1	The former is of equal importance as the latter
3	The former is of higher importance compared to the latter
5	The former is of considerably higher importance compared to the latter
7	The former is of strongly higher importance compared to the latter
9	The former is of extremely higher importance compared to the latter
$2, 4, 6, 8$	The importance lies between two of the aforementioned value
reciprocal	Converse importance as the statements above

Table 5: Instructions to Fill the Table

Some additional rules follow:

$$a_{i\,j} > 0 \qquad a_{i\,j} \cdot a_{j\,i} = 1 \qquad a_{i\,i} = 1 \tag{1}$$

The filled tables are as follow:

①	Inclusivity
②	Gender Equality
③	Safety
④	Sustainability
⑤	Popularity
⑥	Innovation
⑦	Fair Play

Table 6: Factors Considered

	①	②	③	④	⑤	⑥	⑦
①	1	1/5	1/2	1/3	1/8	1/2	1
②	5	1	1	2	1/3	4	5
③	2	1	1	2	1/3	2	4
④	3	1/2	1/2	1	1/5	2	4
⑤	8	3	3	5	1	6	7
⑥	2	1/4	1/2	1/2	1/6	1	2
⑦	1	1/5	1/4	1/4	1/7	1/2	1

Table 7: Filled AHP table

Therefore, the weighting given by AHP is:

①	②	③	④	⑤	⑥	⑦
0.043	0.188	0.147	0.107	0.411	0.065	0.038

Table 8: Weight Computed By AHP

The precision of this result can be expressed by:

$$0 = (A - \lambda I_n) \cdot v \tag{2}$$

$$\mathbf{C.I.} = \frac{\lambda_{\max} - n}{n - 1} \tag{3}$$

$$\mathbf{C.R.} = \frac{\mathbf{C.I.}}{\mathbf{R.I.}} \tag{4}$$

	1	2	3	4	5	6	7	8	9	10
R.I. Value:	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.49	1.51

Table 9: Accepted **R.I.** Values
Formula: $\mathbf{R.I.} = \frac{1}{n}\sum_{\xi=1}^{n} \mathbf{C.I.}_{\xi}$

Where A is the adjusted AHP matrix, I_n is the identity matrix of dimension n, and v is the eigenvector(s). The largest possible λ is denoted as $\lambda_{\max}$, which can be computed by a Python program. For a perfect matrix, $\lambda_{\max} = n$.

C.R. represents the Consistency Ratio, which is determined by the Consistency Index (**C.I.**), a measure of the consistency of the matrix, and Random Consistency Index (**R.I.**), a value influenced by the number of dimensions in the matrix. The accepted value of **R.I.** is shown in table 9. For a precise matrix, the matrix must satisfy **C.R.** < 0.10. The filled matrix in table 7 has a **C.R.** value of 0.024.

4.2 Shannon Entropy Method

The Shannon Entropy method assumes that the more discrete a set of data is, the more information it contains. Examples of this trend are writing systems; A single English letter - which there are 26 of, can arguably express much more information than a Hindu-Arabic numeral - which there are only 10 of. Meanwhile, a Chinese character - which there are thousands of, can express vastly more information than an English letter.

The Shannon Entropy method calculates the weight of each index according to the dispersion of data within each index. In short, the more concentrated a set of data for an index is, the less weight the index is assigned.

To find the weight of each index, the weight of every data entry within each index's data set is first calculated. The formula defining the weight of the j^{th} entry in the i^{th} row (i^{th} index) is as follows:

$$z_{i\,j} = \frac{x_{i\,j}}{\sqrt{\sum_{i=1}^{n} x_{i\,j}^2}} \tag{5}$$

Where n stands for the number of dimensions in the matrix, in this case, 7.

The weight of each data entry within the i^{th} Index's column is then used to calculate the Entropy value of the Index. The Entropy value e_j of each index is a numerical value that represents the amount of disorder in each set of data, which is calculated using the formula:

$$e_j = -k\sum_{i=1}^{n} z_{i\,j}\ln z_{i\,j}\ \text{, where } k = \frac{1}{\ln(n)},\ e_j \geq 0 \tag{6}$$

After obtaining the entropy value of each index, the coefficient of variation g_j of each index is then calculated using the formula:

$$g_j = 1 - e_j \tag{7}$$

Using the coefficient of variation g_j, the final weight of each index can then be calculated by using the formula:

$$w_{\text{Entropy}} = \frac{g_j}{\sum_{j=1}^{m} g_j} \tag{8}$$

Where m stands for the total number of data values the index's data set contains.

After the above calculations, here is the weighting of each index given by the Entropy method (See table 6):

①	②	③	④	⑤	⑥	⑦
0.085	0.032	0.039	0.075	0.679	0.065	0.023

Table 10: Weight Computed By the Entropy Method

4.3 Combining AHP and Entropy Method

To jointly use the AHP and Entropy method, the geometrical average was found using the formula:

Algorithm 11 Pseudocode for Combining AHP and Entropy Method

$$w_{\text{final}ⓘ} \leftarrow \sqrt{w_{\text{AHP}ⓘ} \cdot w_{\text{Entropy}ⓘ}}$$

As a result, the final weighting of each index given by combining AHP and the entropy method is:

①	②	③	④	⑤	⑥	⑦
0.061	0.078	0.076	0.90	0.528	0.065	0.094

Table 11: Weight Computed By Combing AHP and the Entropy Method

5 Results

5.1 Model Output

Based on the weights Table, the index score for each sport is calculated according to the following formula:

Algorithm 12 Pseudocode for Calculating Final Score

$$s_{\text{GlbM}} \leftarrow \begin{array}{l} w_{\text{final}①} \times I_{\text{GInc}} + w_{\text{final}②} \times I_{\text{Gen}} + w_{\text{final}③} \times I_{\text{Safe}} + w_{\text{final}④} \times I_{\text{Sus}} + \\ w_{\text{final}⑤} \times I_{\text{GlbPop}} + w_{\text{final}⑥} \times I_{\text{Inn}} + w_{\text{final}⑦} \times I_{\text{Fair}} \end{array}$$

$$s_{\text{AuM}} \leftarrow \begin{array}{l} w_{\text{final}①} \times I_{\text{GInc}} + w_{\text{final}②} \times I_{\text{Gen}} + w_{\text{final}③} \times I_{\text{Safe}} + w_{\text{final}④} \times I_{\text{Sus}} + \\ w_{\text{final}⑤} \times I_{\text{AuPop}} + w_{\text{final}⑥} \times I_{\text{Inn}} + w_{\text{final}⑦} \times I_{\text{Fair}} \end{array}$$

As the score increases, so does the probability of the SDE already being an Olympic sport.

From the results chart of the Australian model (figure 7) and the General Model (figure 6), we found that there is a score cut-off from 0.387 to 0.529. This range was found to be the margin between unqualified SDEs, to qualified SDEs. The SDEs identified to be underqualified for the Olympic Games, that is, Boxing, Breaking, and Karate, are at the lower part of the general ranking. They have all been planned to be removed from the Olympic Games. Currently, they are not planned to be in the 2028 or 2032 Olympics as well.

SDEs such as Tennis, Volleyball-Indoor, Aquatics-Swimming, and Handball-Indoor are all sports that have continuously been in the Olympics since 1988 or earlier.

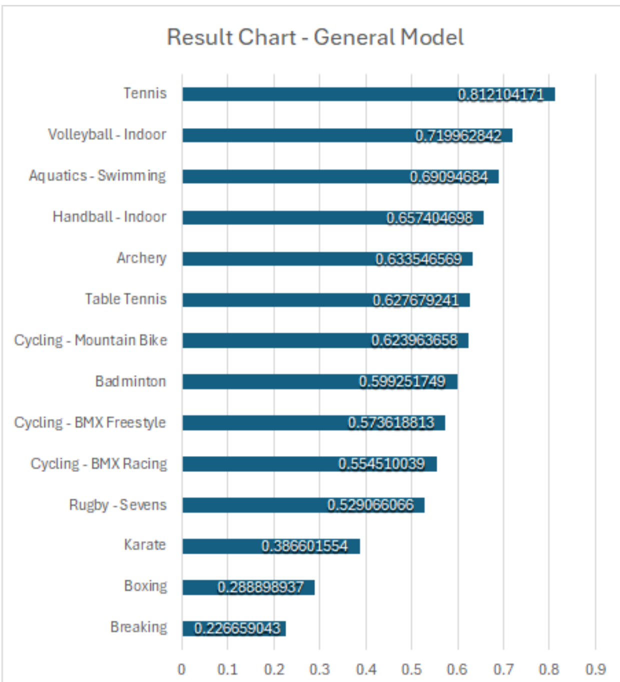


Figure 6: General Model - Selected sports

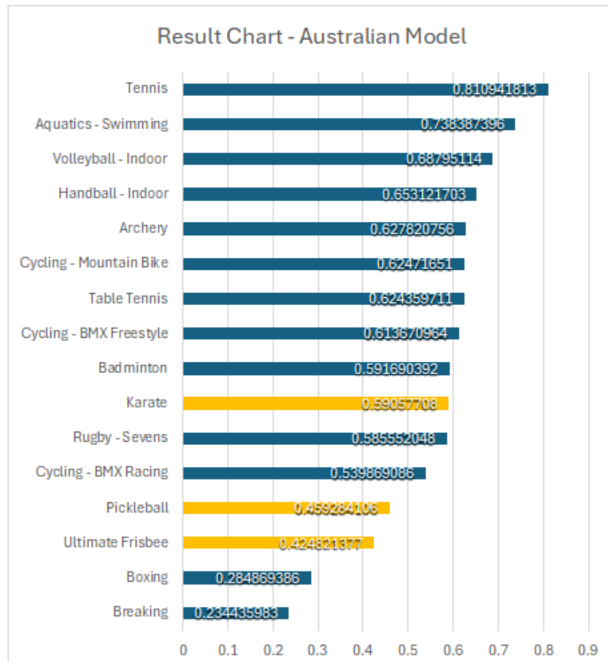


Figure 7: Australian Model w. 3 extra SDEs

5.2 Additional SDEs

Though the model has been proven to work for preexisting SDE's, our next task was to identify and rank 3 additional SDEs for *Brisbane 2032*. By applying the Australian Model on recently removed SDEs as well as recently popular and novel sports, Karate, Pickleball, and Ultimate Frisbee™ were identified as the top 3 candidates, with Karate ranking first in priority followed by Pickleball and Ultimate Frisbee ranking second and third (figure 7).

5.2.1 Karate

For a sport that was removed from the Olympics following *Tokyo 2020* due to lack of entertainment value, Karate has scored extremely well on the Australian Model, with its final score increasing from 0.387 in the General Model to 0.591 on the Australian Model. The reintroduction of this former Olympic sport could help bring larger local audiences, as well as give the former Olympic sport a second chance.

5.2.2 Pickleball

Pickleball is another potential candidate to be added to the Olympics. Scoring a final score of 0.459, this sport has also surpassed the threshold to be added into the Olympics.

Another strong argument for the inclusion of Pickleball in the Olympics is its fast growth. Experiencing a surge in popularity over the past few years according to Google Trends™[10], Pickleball would have the potential to bring in both younger audiences, as well as retain an older fanbase.

Furthermore, the international governing body of Pickleball, the International Pickleball Federation (IPF), was only founded 8 years ago but already had 77 member countries. Though not officially recognized by the IOC yet, we believe that the current patterns of growth as well as the addition of Squash - another popular racket sport, indicate that Pickleball has strong potential to be added to *Brisbane 2032*, or other Olympic games further down the line.

5.2.3 Ultimate Frisbee

Ultimate Frisbee™, with a final score of 0.425 in the General Model, has 107 member countries in its sports governing body: the World Flying Disk Federation (WFDF). It's nature as both a popular, and more notably, it's emphasis on sportsmanship as a self-officiated sport even at elite-level competitions [11] would make Ultimate Frisbee a reasonable candidate for *Brisbane 2032*.

5.2.4 Future Olympics

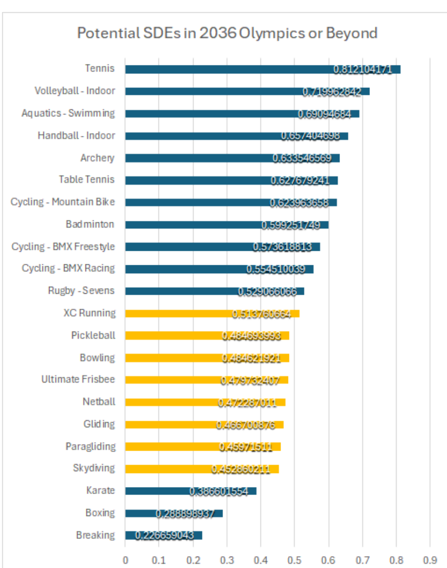


Figure 8: Potential Candidates for Future Olympics

Figure 8 shows more SDEs that could be added to the 2036 Olympics or other Olympic Games further down the line.

This chart is produced using the General Model, and therefore, it does not represent a specific host country and may contain different results compared to figure 7, as well as a lack of country-specific accuracy that may be needed to finalize IOC decisions.

In the figure 8, XC (Cross Country) Running, Pickleball, Bowling, Ultimate Fisbee, Netball, and Air Sports (Gliding, Paragliding, and Skydiving) have all scored high enough to be considered potential candidates for future Olympics. This indicates these SDEs not only have potentials for growth in the future, but also have final scores that are already within the cut-off boundary of 0.387 to 0.529 in figure 7, indicating that they can be categorized as potentially qualifying SDEs.

6 Analysis

6.1 Sensitivity Analysis

The sensitivity of an index represents the magnitude of influence the index has on the final score of an SDE. To test the sensitivity of an index, the weight of the index is either increased or decreased, causing a shift in the weight of all indices. The sensitivity of the index can than be determined by the shift in ranking of all the SDEs after the change in weight.

To estimate the sensitivity of the model as a whole, two indices' sensitivities were calculated: the Popularity Index and the Gender Equity Index.

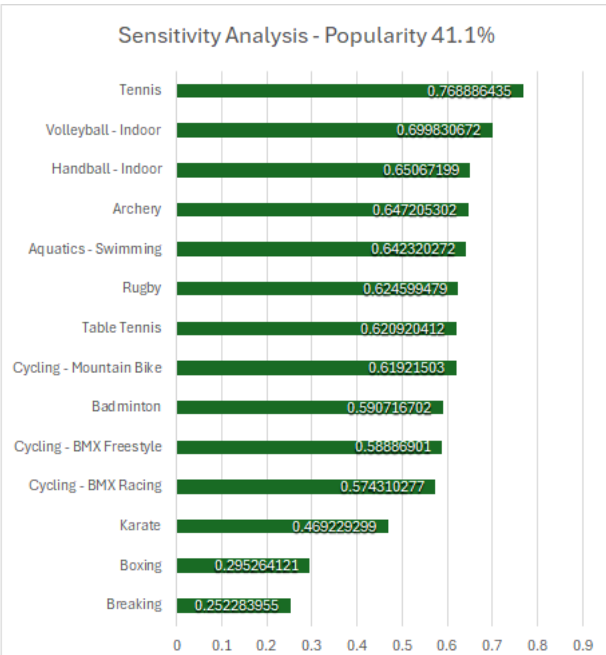


Figure 9: Sensitivity of Popularity

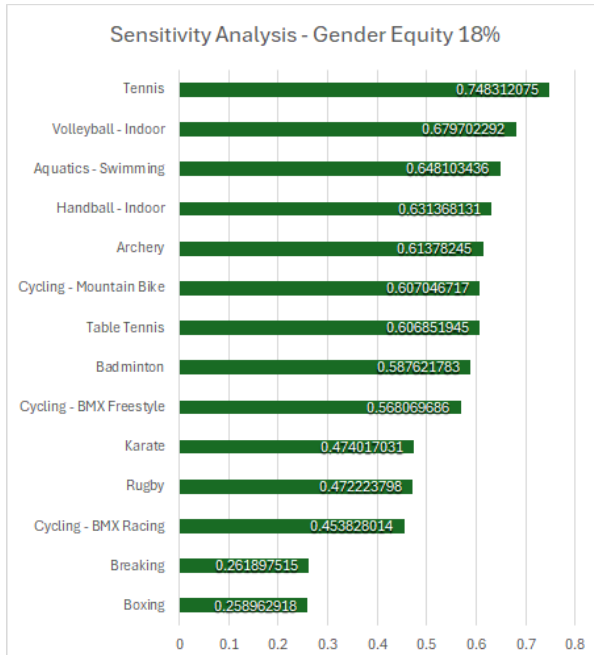


Figure 10: Sensitivity of Gender Equity

The Popularity Index's weight was decreased by 11.7%, which resulted figure 9.

In comparison to the unaltered model, the rankings of SDEs did not undergo significant changes. Only a few SDEs, such as rugby and aquatics - swimming, experienced a shift ranking. Overall, the chart's conclusion whether an SDE qualifies stayed the same, indicating that the popularity index is minimally sensitive.

The Gender Equity Index's weight was increased by 11%. The results are depicted in figure 10.

The resulting graph did not show any significant changes in ranking of the SDEs as well. The two most important indices to the model have proven to have low sensitivity and impact on the model. Therefore, it can be inferred that the model has a low sensitivity overall.

7 Strengths and Weaknesses

7.1 Strengths

1. We used both a subjective evaluation model (AHP) and an objective evaluation model (Shannon Entropy Method), minimizing the amount of error in our combined model.

2. We designed two models for our popularity index, including one that is location-specific and one that is universal. Using either model allows us to rank SDEs for Olympics taking place in a specific country and for multiple Olympics to come.

3. We utilized data from most sports included in *HiMCM_Olympic_Data.xlsx* to construct our Shannon Entropy Method model, ensuring that we have the most accurate weighting for each index.

4. We have more than enough SDEs in our results tables, providing a visual and specific representation of each SDE's placement in the rankings.

5. The results of our model are accurate to the decision by the IOC because the SDEs our model categorized as unqualified are currently not in the Olympics.

6. The weighting for each index determined by the AHP model and Entropy Model has a considerably large difference. This difference further justifies our method of combining models, since the difference indicates inherent biases within both models, and taking a geometrical average balanced the biases.

7.2 Weaknesses

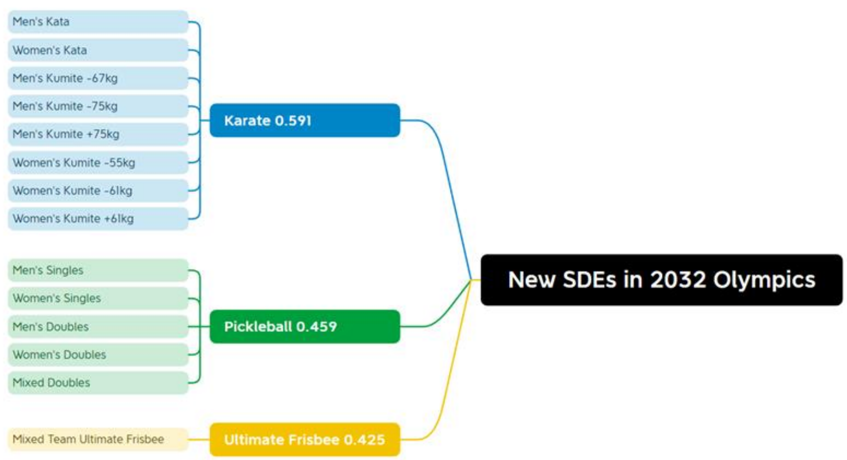
1. Our popularity index is determined by the number of searches of an SDE conducted in English using Google™. However, we did not take into account the variety of dialects in the same language, where different terms are used to address the same SDEs, as well as different search engines being used whose data are not archived in Google Trends, leading to less data collected. This can be improved by including multiple languages in the API and including data from other search engines.

2. Although combined with the Shannon Entropy Method, the AHP models still contain biases nonetheless. If another group attempts to recreate this model, it could potentially yield a different result or even result in failure.

8 Letter to the IOC

Dear International Olympic Committee and to whom it may concern,

As part of the HiMCM Olympics Consulting team, we have been tasked with finding and evaluating aspects of SDEs to determine whether they should be added to the Olympics. To do so, we constructed a mathematical model with both the need for fair, equal, and inclusive sports, as well as popularity and relevancy in mind. For more precise results, we then developed two models for ranking SDEs: A General Model that could be used for ranking SDEs universally or in any Olympics, and a location-specific model that focuses on ranking SDEs specifically for *Brisbane 2032*.
For reference, our results are presented in the diagram below:



To aid in your organization's decision, we have determined the top 3 SDEs, each with their own explanations and scores, that we believe to be the most suitable to add to the 2032 Olympics. The three sports are Karate, Pickleball, and Ultimate Frisbee™.

1. Karate, an SDE that was removed following *Tokyo 2020*, has seen a high score on the Australian-specific Model, and would attract local popularity at *Brisbane 2032*. Although Karate has a low popularity globally, it is more popular in Australia specifically. If for the purpose of the 2032 Summer Olympics only, Karate is qualified to be in the Olympic Games. Australia-Specific Score: **59.1**%, General Score: **38.7**%.

2. Pickleball, an SDE of growing popularity, already has a final score of **45.9**%, which is still expected to grow in the future. Though this sport is not officially recognized yet, we believe that it would see great success in bringing a new generation of viewers to the Olympics. Australian-Specific Score: **45.9**%. General Score: **48.5**%.

3. Ultimate Frisbee has a final score of **42.5**%, emphasizes sportsmanship in its games and is also categorized as a qualified SDE. Notably, it is self-officiated sport, emphasizing sportsmanship. Its introduction would both add more variety to the Olympics as a disc sport, as well as encouraging sportsmanship in other sport. Australian-Specific Score: **42.5**%, General Score: **48.0**%.

Therefore, we recommend Pickleball as the first choice, but Ultimate Frisbee is a close second. Karate is also a contender in *Brisbane 2032*, but for Olympics afterwards, Karate is not as qualified as the previous two.
We sincerely hope these data will be helpful to your organization's decision of what sports to be added and removed in the 2032 Brisbane Summer Olympics.

Thank you for your time,
Team 15848

9 References

[1] International Olympic Committee. *Key milestones in the IOC's history*. URL: https://olympics.com/ioc/history/institutional. Accessed: 2024-11-07.

[2] Rhianna Schmunk. "The Seine is typically filthy. What to know before Olympic swimmers dive in". In: *CBC/Radio-Canada* (July 2024). URL: https://www.cbc.ca/news/world/paris-olympics-seine-water-quality-sewage-1.7265482. Accessed: 2024-11-11.

[3] Caroline Brogan. *Best ways to cut carbon emissions from the cement industry explored*. May 2021. URL: https://www.imperial.ac.uk/news/221654/best-ways-carbon-emissions-from-cement/. Accessed: 2024-11-16.

[4] Statista. *Market share of leading desktop search engines worldwide from January 2015 to January 2024*. Feb. 2024. URL: https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/. Accessed: 2024-11-18.

[5] International Olympic Committee. *1913 : First public presentation of the Five Rings Symbol*. URL: https://olympics.com/ioc/1913-first-public-presentation-of-the-five-rings-symbol. Accessed: 2024-11-12.

[6] International Table Tennis Federation. *Member Associations & Continental Federations*. URL: https://www.ittf.com/mas-and-continents/. Accessed: 2024-11-08.

[7] International Olympic Committee. *2021-2024 Gender Equality and Inclusion Objectives*. URL: https://olympics.com/ioc/gender-equality/objectives. Accessed: 2024-11-12.

[8] Q2032. *Venues and villages*. URL: https://q2032.au/plans/venues-and-villages. Accessed: 2024-11-16.

[9] Luis G. Vargas. "The Legacy of the Analytic Hierarchy/Network Process". In: *International Journal of the Analytic Hierarchy Process* 9.3 (2017). DOI: https://doi.org/10.13033/ijahp.v9i3.541.

[10] Google. *pickleball*. Jan. 2024. URL: https://trends.google.ca/trends/explore?date=2014-01-01%202023-12-31&q=pickleball&hl=en-us. Accessed: 2024-11-18.

[11] USA Ultimate. *governance*. URL: https://usaultimate.org/about/. Accessed: 2024-11-18.

10 Appendix

```python
import numpy as np
from pytrends.request import TrendReq
import matplotlib.pyplot as plt
from time import sleep
import pandas as pd
request = TrendReq()

queries = ["Queries to search up"]
names=["Labels for the legend"]

# Word to compare everything against
standard = "Artistic Swimming"
max_val = 0
points = []
df = []
for i in range(len(queries)):
    if queries[0] == standard:
      query = [queries[0]]
    else:
      query = [queries[0], standard]
    request.build_payload(kw_list=query, timeframe="2004-12-31 2024-11-10"
    )
    df = request.interest_over_time()
    sleep(3)
    y = df[queries[0]]
    x = [df.index[i] for i in range(len(y))]
    max_of_current = max(df[queries[0]])
    if max_of_current > max_val:
        max_val = int(max_of_current)
    points.append([x, y])
    queries.pop(0)

# Graphs points. I just saved to an excel spreadsheet
points = np.array(points)
for i in range(len(points)):
    points[i][1] *= 100/max_val
    plt.plot(points[i][0], points[i][1], label=names[i])
plt.ylabel("% Popularity")
plt.xlabel("Time")
plt.legend()
plt.show()
```