Problem Chosen	2023 MCM/ICM Summary Sheet	Team Control Number
Z		2333325

Measurement of Olympic Impact and Evaluation of Olympic Strategies and Policies

Summary

Honored as the world's most influential sports event, the Olympic Games have been held 33 times to date, bringing both burden and benefits to the host country. Facing the opportunities and challenges brought by hosting the Olympics, it is important to establish a mathematical model to measure the impact of the Olympics and discuss the feasibility of choices, strategies, and policies. We presented our work in three phases.

Firstly, based on extensive literature research, we established the **Olympic Games impact evaluation model** from six aspects: **economic, land use, human satisfaction, travel, opportunity for future improvements, and prestige**. In this evaluation model, we selected seven indicators for quantification. To objectively measure the contribution of different indicators to the influence, we used the **Entropy Weight Method (EWM)** to determine the weights of different indicators. To quantify the impact of a given Olympic Games data on relevant indicators, we used the **weighted TOPSIS method** to calculate the score of the Olympic Games and quantify its impact.

Secondly, we discussed the **feasibility** of some potential policies and strategies, and established a reasonable Olympic Games preparation **timeline**. For the feasibility of policies and strategies, we explored the pros and cons of some selected policies and strategies from the perspectives of economic benefits, social acceptance and risks, thus making judgments on the feasibility of these policies. We also established a reasonable Olympic timeline, provided preparation time nodes for important tasks, and discussed the significance and necessity of considering such factors.

Finally, based on the inputs and outputs of past Olympic Games in different aspects, we used the method of **Multiple Linear Regression** to establish the **impact prediction model of Olympic options**. In this model, the inputs of the Olympic Games were mainly divided into five aspects: facility and venue investment, land use, sponsorship quantity, security investment, and human investment; while the outputs included three aspects: expected revenue, cultural influence, and building reuse value. Before model building, we conducted **outlier and correlation tests** on the data. Correlation tests were also performed to check whether our chosen explanatory variables were appropriate for the model. After we built multiple linear regression model, we conducted a **goodness-of-fit test** on this model to ensure its effectiveness.

Keywords: Olympic Games impact; Weighted TOPSIS; Feasibility; Timeline; Multiple Linear Regression.

Contents

1 Introduction

1.1 Problem Background

The hosting of the Olympic Games would bring many long-term or short-term effects to the country, such as economic benefits, cultural influence and national prestige. Nowadays, the International Olympic Committee (IOC) is facing a decline in the number of bids for both the Summer and Winter Games, largely because of the cost pressures that come with hosting the Games. In order to study the impact of the Olympic Games more clearly and provide better suggestions on hosting the Olympics Games, we developed metrics to measure the impact of the Olympic Games. Meanwhile, we worked to find reasonable policies and strategies to support the hosting of the Olympic Games.

1.2 Restatement of the Problem

Considering the background information and restricted conditions identified in the problem statement, we needed to solve the following problems:

- Determine metrics for the impacts of hosting the Olympic Games
- Propose policies and strategies, and consider the feasibility of them
- Discuss the timeline for preparing for the Olympics and the significance of the implementation of the timeline.
- Evaluate the impact of potential options and strategies

1.3 Our Work

To begin with, we collected the data of the Olympic Games in the past years, including the expenditure and income of the Olympic Games, the land area, the number of viewers and other 10 indicators about the 17 Olympic Games. Using the TOPSIS method with entropy weight, we established the Olympic Games influence evaluation model with 6 influencing factors. Then we put forward some possible policies and strategies and discuss their feasibility in a reasonable way. After that, we gave a timeline for implementing the Olympic Games and gave its significance. Finally, we used the method of Multiple Linear Regression to obtain three regression equations on different measures to quantify the impact of the Olympic Games and build the impact prediction model of Olympic options. Our work can be shown as follows:

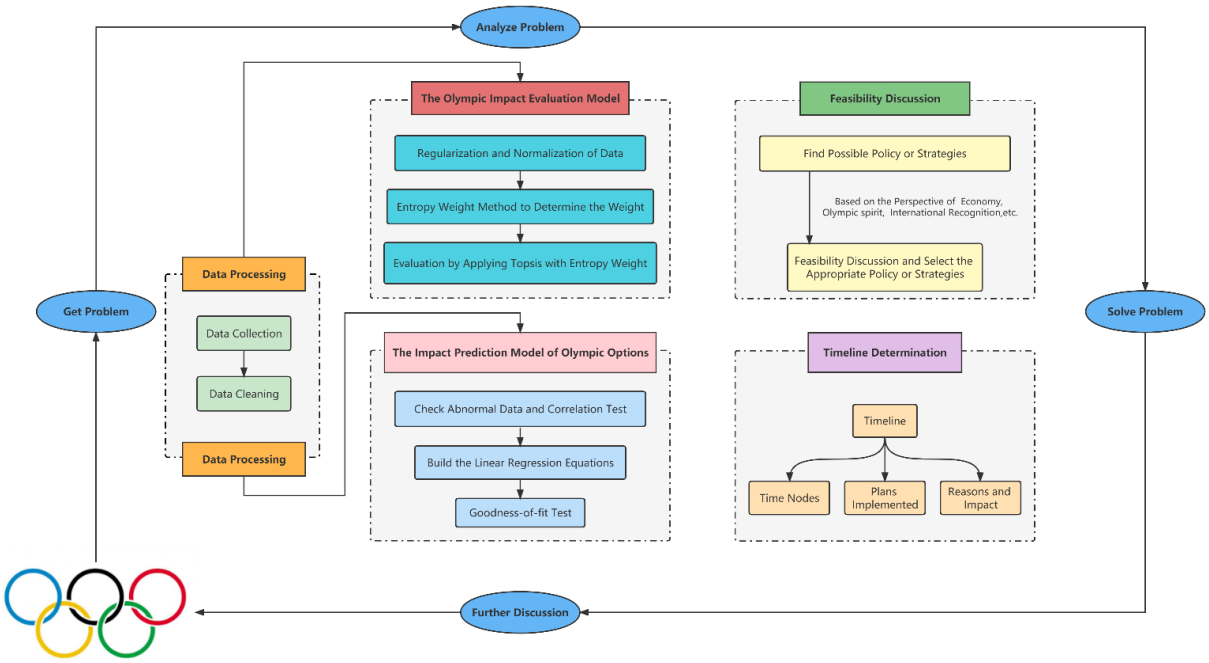

Figure 1. Our Works.

2 Assumptions and Justifications

Assumption 1: The data we collect is accurate and rigorous.
Justification: We collected relevant data on a large number of official websites, which can generally be considered accurate and rigorous. On this basis, we can establish a reasonable Olympic impact evaluation model and an Olympic plan impact prediction model.

Assumption 2: The indicators we have chosen can effectively reflect the impact of the Olympics in one aspect that we want to discuss.
Justification: There are many indicators that can reflect the influence of the Olympic Games to some extent, and there are also interactions between different indicators. It is not realistic to list all the relevant indicators in detail, and it will also introduce bias into the model. Therefore, some indicators that mainly reflect the major aspects of influence are chosen.

Assumption 3: Assume that only the indicators and data related to the Olympic Games after 1992 are reasonable.
Justification: 1992 was the beginning of the "professionalization" of the Olympic Games, which were more modern and professionalized; Meanwhile, the 1992 and subsequent Olympic Games were relatively close in time to the current Olympic Games, therefore, they had more reference value and significance when we compared the influence of each Olympic Games.

Assumption 4: Suppose the Olympic host authorities took a positive attitude towards hosting the Olympics.
Justification: In order to increase the positive effects of the Olympic Games and reduce the negative effects, the host authorities often takes a positive attitude.

3 Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1: Table of notations

Symbol	Description
z	The result after normalization
w_j	The weight of the ith indicator determined by the entropy weight method
S_i	Score of the ith Olympic Games before applying TOPSIS.
$\hat{S}_i$	Score of the ith Olympic Games after applying TOPSIS.
R^2	The goodness-of-fit of the impact prediction model for the Olympic scenario
$y_{i,j}$	The actual value of the ith Olympic Games at the jth factor
$e_{i,j}$	The residuals of the ith Olympic Games on the jth factor, $e_{i,j} = \hat{y}_{i,j} - y_{i,j}$

4 The Olympic Games Impact Evaluation Model

4.1 Data Description

The Olympic Games, as one of the most important events in the world and even the whole mankind, has a multidimensional and comprehensive influence on every field of society. Therefore, when constructing The Olympic Games Impact Evaluation Model, we needed to seriously consider all aspects of the Olympic Games influencing factors and chose appropriate indicators to measure these impacts.

In the establishment of the impact evaluation model below, we collected data on 10 different indicators of 6 aspects of the 17 Olympic Games from 1992 to 2022.

(1) Economy

When measuring the economic benefits brought by the Olympic Games, we mainly considered the economic return brought by hosting the Olympic Games itself (i.e. “relative payoff”) and “the percentage of GDP growth eight years after the Olympic Games”. Then we took these two indicators as the main indicators to measure the economic impact brought by the Olympic Games. The formula for calculating relative payoff is as follows:

$$\text{relative payoff} = \frac{\text{Olympic revenue} - \text{Olympic expenditure}}{\text{GDP}}$$

Therefore, in terms of economy, we needed to collect data of the following indicators.

Table 2: Table of Economic Indicators to be Collected.

Index	Unit	Remark
Olympic revenue	U.S. dollar $	/
Olympic expenditure	U.S. dollar $	/

GDP	U.S. dollar $	/
the percentage of GDP growth eight years after the Olympics	%	A benefit-type indicator
relative payoff	%	A benefit-type indicator

(2) Land Use

Large amounts of land were required to host the Olympics. Since the heavy consumption of land was essentially a consumption of resources, we measured the negative impact of the Olympic Games on land use in terms of "land occupied area". Obviously, "land occupied area" was a cost-type indicator.

(3) People's Satisfaction

The positive impact of the Olympics was also reflected in people's satisfaction with the Games. This can be reflected in the indicator " viewership". Obviously, it is a benefit-type indicator.

(4) Travel

The location of the Olympic Games tended to attract more tourists. When considering the impact of the Olympic Games on tourism, we could use the benefit-type indicator "growth rate of tourism of the host city compared to the previous year".

(5) Opportunities for Future Improvement

There was still a possibility that the Olympic site could be used in the future. According to the relevant information on the Internet, we could understand the use of each Olympic venue. After that, we gave rates according to their utilization, the lowest point is 1, the highest point is 5. We could measure their opportunities of future improvement through "utilization scores" (benefit-type indicator).

(6) Prestige

The Olympics brought hot spots and prestige to host countries and host cities. It was natural to use “the number of search terms” as a benefit-type indicator to measure the prestige impact of the Olympics on the host countries and host cities.

4.2 Data Processing

The data collected were shown in Appendix 1. Next, we would focus on the preliminary processing of the data.

4.2.1 Missing Value Processing

The process of collecting data from multiple countries and different time points could make it difficult to ensure that all data is fully complete. However, it was essential to obtain available data. Therefore, it was necessary to handle missing data properly to improve the accuracy and validity of our model. We applied two methods for processing missing data.

(1) Same-class Mean Interpolation Method

This method involved predicting the category of a missing value using a cluster analysis model, and then replacing the missing value with the mean value of that category.

(2) Mean Interpolation Method:

This method involved replacing missing interval data with the mean value of the corresponding category. For non-numerical data, the mode value was used for imputation.

4.2.2 Data Normalization

The indicators we mentioned before could be divided into cost-type indicators and benefit-type indicators. For these two different indicators, we used different normalization methods.

For cost-type indicators,

$$z = \frac{max - x}{\max - min},$$

For benefit-type indicators,

$$z = \frac{x - min}{\max - min},$$

where "*max*" means the maximum value of this indicator, "*min*" means the minimum value of this indicator.

After finishing the data normalization, we could obtain a 18×7 normalized matrix, which could be useful later

$$Z = \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,7} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,7} \\ \vdots & \vdots & \ddots & \vdots \\ z_{18,1} & z_{18,2} & \cdots & z_{18,7} \end{bmatrix}.$$

4.3 Entropy Weight Method to Determine the Weight

The **Entropy Weight Method (EWM)** is a multi-criteria decision-making technique that calculates the weight of each criterion based on the concept of entropy. The method involves calculating the entropy of each criterion and determining the weight of each criterion by comparing the individual criterion entropy with the total entropy of all criteria. The higher the entropy of a criterion, the lower its weight, and vice versa.

Firstly, we calculated the weight of the element in the ith row of the jth column in Z.

$$f_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}}.$$

According to the notions of self-information and entropy in information theory, the information entropy e_j of each evaluation index in calculation was obtained, which is

$$e_j = \ln(n)^{-1} \sum_{i=1}^{n} f_{ij} \ln(f_{ij}).$$

Based on the information entropy, we would further calculate the weight of each indicator according to the previous definitions (the weight of 7 indicators).

$$w_j = \frac{1 - e_j}{n - \sum_j e_j}, \qquad j = 1,2, \dots, n.$$

The weight of each index was now obtained, which was shown in the figure below:

Table 3: Index Weight of the Influence Evaluation Model.

index	weight
Relative payoff (w_1)	0.2417
The percentage of GDP growth(w_2)	0.3933
Land occupied area(w_3)	0.1355
Viewership (w_4)	0.0565
Growth rate of tourism(w_5)	0.0171
Utilization scores (w_6)	0.0640
The number of such terms (w_7)	0.0925

4.4 TOPSIS Method with Entropy Weight

The TOPSIS method was a commonly used multi-criteria decision analysis method for evaluating the superiority and inferiority of decision alternatives, which treated decision alternatives as combinations of multiple evaluation criteria.

To perform TOPSIS with Entropy Weight, we needed to do the following steps and use the normalized matrix to calculate some important intermediate variables.

$$Z = \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,7} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,7} \\ \vdots & \vdots & \ddots & \vdots \\ z_{18,1} & z_{18,2} & \cdots & z_{18,7} \end{bmatrix}$$

(1) Find the maximum vector of each index,

$Z^+ = (Z_1^+, Z_2^+, \dots, Z_7^+)$, where $Z_j^+ = \max\{z_{1,j}, z_{2,j}, \dots, z_{18,j}\}$

(2) Find the maximum vector of each index,

$Z^- = (Z_1^-, Z_2^-, \dots, Z_7^-)$, where $Z^- = \min\{z_{1,j}, z_{2,j}, \dots, z_{18,j}\}$

(3) Find the distances between each evaluation target and the maximum vector

$D_i^+ = \sqrt{\sum_{j=1}^m w_j \times (Z_j^+ - z_{i,j})^2}$.

(4) Find the distances between each evaluation target and the minimum vector

$D_i^- = \sqrt{\sum_{j=1}^m w_j \times (Z_j^- - z_{i,j})^2}$.

Then, we calculated scores of each evaluation by measuring how close they were to the maximum and minimum vectors, that was

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-}.$$

Since our aim was to evaluate the impacts of hosting the games, we could calculate the corresponding S_{18} to see how much the impact of the Olympics Games is if we were given

the data of the indicators in our model.

5 Feasibility and timeline to implement

Firstly, it was important to acknowledge that not all strategies put forward were advantageous. For instance, some advocated dividing Olympic sports into four groups instead of two. However, after analysis, this paper deems this approach unfeasible and reasons are as follows:

To begin with, it would diminish the uniqueness and allure of the Olympic Games. The Olympics are a symbolic international sports event that occurs every four years, with its cyclical and grand nature giving it exceptional significance. Frequent hosting of the Olympics, as proposed, would reduce its distinctiveness and may result in decreased public interest. Furthermore, hosting the Olympics has a substantial environmental impact, including increased energy consumption, traffic congestion, and waste disposal. As such, frequent hosting of the Olympics would inevitably impose a greater burden on the environment, potentially leading to environmental pollution and ecological damage. Lastly, frequent hosting of the Olympics results in a continuous loss of funds for maintenance, adding to the financial strain of the host country.

In light of these concerns, this paper identified three alternative strategies that were viewed as more advantageous and feasible.

5.1 Feasibility Analysis

5.1.1 The Potential Strategies to be Applied

(1) Multiple Host Cities

Description: The Olympics have more than one city to host the Olympic Games.

Assigning more than one city to host the Olympics is blessed with lots of advantages, including sharing cost, increasing visibility, spreading risk, promoting regional cooperation, and enriching the competition content.

[1] Cost Sharing: A large-scale international sports event requires a lot of financial expenditure, including venue construction, security measures, transportation and other aspects. If it is held in multiple cities or regions, these costs can be spread across different places, reducing the burden on each city.

[2] Increase Visibility: Hosting the Olympics is an opportunity for countries and cities to showcase their image and strength. Multi-city hosting allows more cities to receive attention and recognition, thus increasing the popularity and influence of them. What's more, the increase in popularity of the host city makes for promoting the economic and social development of them.

[3] Spread Risk: Risks always exists associated with hosting the Olympics, such as security issues, weather and other unpredictable factors. If Olympic games are held in multiple cities, these risks can be spread and the impact on the entire event can be reduced compared to single-

city hosting.

[4] Promote Regional Cooperation: Hosting the Olympic Games in multiple cities requires coordination and cooperation between different cities, which can promote exchanges and interaction between cities and strengthen interregional cooperation and development.

[5] Enrich the Competition Content: Hosting the Olympic Games in multiple cities can distribute the competition events among different cities, hence enriching the competition content and improving the audience's experience when viewing the matches, thus attracting more audience and media attention.

(2) Co-hosting of the Olympic Games by Many Countries

Description: Disperse the Olympic Games across multiple countries and regions to improve efficiency and economic benefits.

It is widely acknowledged that co-hosting of the Olympic Games by several countries has benefits in various aspects, like sharing the financial pressure, improving international cooperation, promoting cultural exchanges, and enhancing the influence of the Olympic Games.

[1] Share Financial Pressure: Hosting the Olympic Games requires a lot of capital investment, which puts great pressure on a country's finances. Co-hosting can reduce the burden on each country and lower the total cost of the Games compared to single-hosting.

[2] Improve International Cooperation: Close cooperation between host countries is always required to successfully host the Olympic Games, which can promote mutual trust and cooperation between countries and enhance their international image and status. Therefore, multinational co-hosting of the Olympic Games can undoubtedly promote international cooperation.

[3] Promote Cultural Exchanges: Different countries and regions have great cultural differences, and co-hosting can provide a broader platform for cultural display, attract more countries and regions to participate, and thus promoting cultural exchanges and understanding.

[4] Enhance the Influence of the Olympic Games: Co-hosting can expand the Games to a wider range of area, attracting more spectators and media attention. Therefore, Multinational co-hosting of the Olympic Games can increase the visibility and influence of the Games.

(3) Combine Summer Olympics and Winter Olympics

Description: Combine the summer Olympics and the winter Olympics to be the Spring & Summer & Winter Olympics.

Although there is no precedent, changing the Olympics to Spring & Summer & Winter Olympics is likely provide benefits like improving resource utilization efficiency, reducing environmental impact, increasing diversity and interest, and increasing participation and influence.

[1] Improve Resource Efficiency: Spring & Summer & Winter Olympic Games co-hosting can make more efficient use of venues, equipment and human resources. The construction of Olympic venues and equipment requires huge investments and has a limited lifespan. By combining the Summer and Winter Olympic Games to host only one Spring & Summer & Winter Olympic Games, these venues and equipment can be used more efficiently, reducing

investment costs and resource waste.

[2] Reduce Environmental Impact: In contrast to the idea of splitting the Olympic sports into four groups instead of two groups mentioned at the beginning, joint hosting of the Spring & Summer & Winter Olympic Games can reduce the environmental impact of the Olympic Games, reduce energy consumption and carbon emissions, and reduce waste generation and disposal.

[3] Increase Diversity and Interest: Olympic events vary greatly from season to season. Therefore, Combining Summer Olympics and Winter Olympics can bring these events together, hence increasing the variety of the competition. Moreover, since the combination of the two Olympics can bring more interest into the game, more viewers and media attention are likely to be attracted.

[4] Increase Participation and Influence: Some countries and regions may not be able to participate in a particular season of the Olympic Games due to climate and season, but the co-hosting of the Spring & Summer & Winter Olympics can provide them with a wider range of opportunities to participate, thus increasing the participation and influence of the Olympic Games.

5.1.2 The Potential Policies to be Applied

In this section, we mainly present the potential policies that we consider as feasible.

(1) Take Advantage of Existing Venues.

Corresponding factors: **Opportunities of improvement in the future & Economy.**

Utilizing existing venues and facilities brings about the following advantages.

[1] Increase Historical Value: Some venues of historical value were fully utilized and displayed in the Olympic Games, they became part of the local cultural heritage and attracted more visitors.

[2] Improve Organizational Skills: Cities and organizers who have successfully hosted Olympic Games before have accumulated a wealth of experience to better respond to issues and challenges, improve organizational skills and efficiency.

[3] Save Costs: Utilizing existing venues can save huge costs in building new venues, as well as delays and additional costs that may occur during the construction of new venues.

(2) Adjust the Scale of the Olympic Games Combined with National Conditions.

Corresponding factors: **Land use & Economy.**

Occasions exists when the host countries have special situations, for instance, the land resources are relatively scarce. Other situations may include economic strength, the existing infrastructure and the population and so on. Should these happen, the above policies could more or less help them out because of the following reasons.

[1] Improve Resource Efficiency: Limited land resource calls for limited scale, which in tern encourages the improvement in resource efficiency. The host country tends to better take advantage of the resource in order not to reduce the quality of the Olympics.

[2] Save Costs: Once the scale is reduced due to various reasons, the construction costs, operating costs, environmental costs will all be reduced to some extent.

(3) Stick to the Principle of Sustainable Development.

Corresponding factor: **Environment & Prestige of the host country.**

Since the Olympics event only lasts about half a month officially, host countries should attach great importance to sustainable development. Actions can be taken through the use of renewable energy, reducing the use of plastic and paper, encouraging public transport, and so on.

These actions may also benefit the country and the Olympics in several aspects.

[1] Reduce Environmental Pollution and Save Resources: The principle of sustainable development focuses on environmental protection and resource utilization, which can reduce the pollution and damage of the Olympic Games to the environment while saving the resources, hence reducing the negative impact on the ecological environment.

[2] Improve the National Image: Adhering to the principle of sustainable development can improve the national image of the host country and city by showing an image of responsibility, trustworthiness and a good social image.

5.2. Timeline to Implement

A good Olympic implementation timeline can help smooth the planning of the Olympics and hopefully make the Games more influential. Here is our timeline of the Olympics implementation.

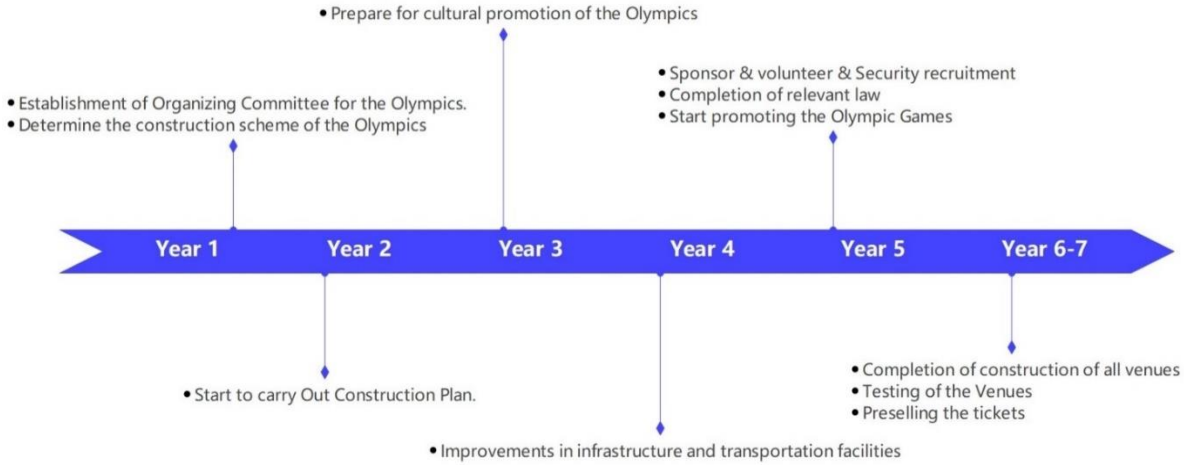

Figure 2. Timeline of the Olympics implementation.

Year 1

- **Establishment of Organizing Committee for the Olympics.**

  The Olympic Organizing Committee is important: its role is to carry out the spirit of the government's directive on the hosting of the Olympic Games; To fulfil its responsibilities and obligations under the Olympic Charter and the Host City Contract under the guidance of the IOC and its Coordinating Commission; To help independently and efficiently carry out the preparatory and hosting work of the Olympic Games.

- **Determine the Construction Scheme of the Olympics**

  Olympic architecture also plays an indelible role in the promotion of urban culture, and is an intuitive perception tool to reflect the development level of a country. We need to

consider the building proposal comprehensively. The development of the building plan includes consideration of building area, site budget, project time, etc., and needs to consider a variety of design concepts, such as environmental protection, energy saving, art design, Olympic spirit, etc.

Year 2

- **Start to Carry Out Construction Plan.**

With a construction scheme in place, it is necessary to start construction for the Olympics the following year.

Year 3

- **Prepare for Cultural Promotion of the Olympics**

Starting from the third year, we are supposed to prepare for the cultural promotion of the Olympic Games, which includes proposing the Olympic ideas and themes, designing the Olympic slogan, making the Olympic logo and so on. The cultural promotion of the Olympic Games should pay attention to innovation and The Times, and combine with the Olympic spirit to enhance the cultural influence of the Olympic Games.

Year 4

- **Improvements in Infrastructure and Transportation Facilities**

Infrastructure and transportation have an important impact on the Olympic Games. Infrastructure is the name card of a country. Improving infrastructure is conducive to improving the national image, facilitating people's lives and realizing sustainable development. Before the Olympics, it is common practice to upgrade transportation facilities, improve public services and use new technologies to improve service quality. This helps to improve citizen and tourist satisfaction, and to some extent also stimulates economic development.

Year 5

- **Sponsor & Volunteer & Security Recruitment**

The cooperation between the Olympic Games and sponsors can mainly reduce the burden of costs. Generally speaking, the admission of sponsorship should consider the sponsor's qualification factor, guarantee factor, quotation factor and brand factor. The sponsor must be a powerful enterprise and a leading enterprise in the industry; There are good prospects for development and sufficient funds to pay for sponsorship. Secondly, sponsors are required to provide adequate, advanced and reliable products, technologies or services for the successful hosting of the Games. In addition, the sponsorship price quoted by the enterprise is one of the most important factors to consider when choosing a sponsoring enterprise. Brand factors are also taken into account. The company has a good social image and corporate reputation, the company's brand and image in line with the Olympic ideal and the concept of the Beijing Olympic Games, products in line with environmental protection standards.

The Olympics need to recruit a large number of volunteers to keep the Games running. Organizers of the Olympic Games need to consider volunteer work content, working hours, recruitment requirements and other aspects to ensure the orderly operation of volunteer work.

The security plan for the Olympics is also important. The host party of the Olympic Games needs to check the potential security problems during the Olympic Games, strengthen the Olympic supervision (such as crimes, possible incidents), and fully arrange the work of the security team.

- **Completion of Relevant Law**

  The purpose of establishing relevant laws for the Olympic Games is to regulate people's behavior under the premise of adapting to the circumstances of the Olympic Games, ensure the smooth running of the Olympic Games and safeguard the dignity and rights of the Olympic Games.

- **Start Promoting the Olympic Games**

  Promoting the Olympics makes for increasing its influence in people's minds. Moreover, an appropriate timing of promotion of the Olympics can adjust the interest of the people in Olympics at the maximum when it officially starts.

Year 6-7

- **Completion of Construction of all Venues**

  All venues should be ready for test by this time.

- **Testing of the Venues**

  Testing of the venues increases the safety rates of the official games.

- **Preselling the Tickets**

  Preselling the tickets increases the popularity of the event and reduces the potential crowding in the official ticket selling in the future.

6 The Impact Prediction Model of Olympic Options

The Olympic Games are one of the largest and most prestigious international sporting events, attracting millions of viewers and generating billions of dollars in revenue. Host cities invest significant resources into hosting the Games, with the expectation of reaping economic benefits and boosting their global image. However, the decision to host the Olympics is never free of risk, as the costs and impacts of hosting can vary widely depending on a range of factors, including the city's infrastructure, the size and scope of the event, and the global economic and political environment. To help assess the potential impacts of hosting the Olympics, researchers have developed various models and methods, including the impact prediction model of Olympic options. This model aimed to provide decision-makers with a quantitative tool for evaluating the economic and social impacts of hosting the Games, as well as for identifying and mitigating potential risks and challenges. In this section, we would explore the key features and applications of the impact prediction model of Olympic options, as well as its strengths and limitations.

6.1 Data Description and Cleaning

We use site investment, land area, number of sponsors, security investment and manpower input as regression independent variables, and Olympic revenue, cultural influence and building reuse value as regression dependent variables. Among them, we used the search index as an indicator of cultural influence and the growth rate of travel compared to the previous year as an indicator of the reuse value of buildings.

To preprocess the dataset on Olympic Games, we followed a series of steps to clean and prepare the data for analysis. In this process, we encountered some data points that were unreasonable or missing.

Firstly, we examined the dataset for any missing values. We found that some of the data points were missing, particularly the data for certain variables such as the number of sponsors for the Olympic Games. We tried to retrieve this missing data from official Olympic reports and statistical sources, but unfortunately, we were unable to find some of the missing data. As a consequence, we decided to remove the missing data points from our analysis.

For missing values, one approach was to impute the missing values with the mean or median value of the dataset. It was a simple method that can be used when the number of missing values is small and the missing values are randomly distributed. The formula for calculating the mean was:

$$mean = \frac{\sum_{i=1}^{n} Value_i}{number\ of\ values}.$$

We could replace the missing values with the mean value of the variable. For example, if we had missing values for the number of sponsors, we could calculate the mean value of the sponsors for the available data and use this value to replace the missing values. However, if the number of missing values was large or the missing values are not randomly distributed, other imputation methods might be necessary.

6.2 Outlier Test and Correlation Test

6.2.1 Outlier Test

Next, we identified outliers and erroneous data points. We used various statistical measures such as mean, median, and standard deviation to identify any data points that were significantly different from the rest of the dataset. After identifying the outliers, we evaluated the reasonability of these data points and either removed them or replaced them with more reasonable values.

To identify outliers and erroneous data points, we used various statistical measures such as the Z-score and the interquartile range (IQR). The Z-score measures how many standard deviations a data point is away from the mean. Any data point with a Z-score greater than 3 or less than -3 is considered an outlier. The IQR is the range between the 75th percentile and the 25th percentile of the data. Any data point that falls outside the range of 1.5 times the IQR is

considered an outlier. The formula for the Z-score is:

$$Z = \frac{x - x_{mean}}{\sigma}$$

where x is the data point, mean is the mean value of the dataset, and standard deviation is the standard deviation of the dataset. The formula for the IQR is:

$$IQR = Q3 - Q1.$$

where $Q3$ is the 75th percentile of the data and $Q1$ is the 25th percentile of the data. We also conducted data validation to check for data integrity issues, such as inconsistent values between different variables. For example, we checked whether the land area and the amount of investment in the venues were consistent with each other, and whether the number of volunteers and the amount of investment in human resources were also consistent.

In summary, our data preprocessing involved identifying missing data, removing outliers and erroneous data points, validating data integrity, and standardizing the data. Despite the challenges we faced, we ensured that the data was as accurate and reliable as possible to enable meaningful analysis.

6.2.2 Correlation Test

When building a multiple linear regression model, it is important to determine whether there is a linear relationship between the independent variables and the independent variables. One way to do this is by using correlation analysis, which involves calculating the correlation coefficient ($r_{i,j}$) between the independent variables X_i and X_j. The correlation coefficient ranges from -1 to 1, with values of -1 indicating a perfect negative correlation, 0 indicating no correlation, and 1 indicating a perfect positive correlation. The formula for the correlation coefficient is:

$$r_{i,j} = \frac{\sum_{k=1}^{n}(X_i^{(k)} - \overline{X_i})(X_j^{(k)} - \overline{X_j})}{\sqrt{\sum_{l=1}^{n}(X_i^{(l)} - \overline{X_i})^2}\sqrt{\sum_{m=1}^{n}(X_j^{(m)} - \overline{X_j})^2}}$$

Thus, we could obtain the following correlation coefficient matrix,

$$r = \begin{bmatrix} 1 & 0.20 & -0.14 & -0.15 & 0.37 \\ 0.20 & 1 & 0.12 & 0.08 & -0.13 \\ -0.14 & 0.12 & 1 & 0.65 & -0.43 \\ -0.15 & 0.08 & 0.65 & 1 & 0.0009 \\ 0.37 & -0.13 & 0.0009 & -0.43 & 1 \end{bmatrix}$$

It's easy to see that all correlation coefficient except for the diagonal is not close to 1 or -1, which means that there is no significant linearity between the independent variables.

6.3 Multiple Linear Regression

To establish three multiple linear regression models between the five explanatory variables (Site investment, Floor area, Number of sponsors, Security investment, Human input) and each

of the three response variables (Predicted income), we first needed to define the mathematical formula for multiple linear regression.

The formula for multiple linear regression could be expressed as follows:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \varepsilon.$$

where Y is the response variable (predicted income), x_1, x_2, ..., x_n are the explanatory variables (Site investment, Floor area, Number of sponsors, Security investment, Human input), β_0 is the intercept, β_1, β_2, ..., β_n are the regression coefficients, and ε is the error term.

Note that this implies:

$$Y_i | X_i = x_i \sim N\left(\beta_0 + \sum_{i=1}^{n} \beta_i x_i, \sigma^2\right).$$

Therefore,

$$E_{f_{Y|X}}[Y_i | X_i = x_i] = \beta_0 + \sum_{i=1}^{n} \beta_i x_i.$$

In vector notation, this model can be re-written $Y_i = x_i^T \beta + \varepsilon_i$, where $x_i = (1, x_{i1}, x_{i2}, \ldots, x_{in})^T$ and thus for vector $\mathrm{Y} = (1, Y_1, \ldots, Y_n)^T$ we have

$$Y = X\beta + \varepsilon.$$

For projected income, the result went as follows:

Table 4: Projected Income Regression Coefficient Table.

β_0	β_1	β_2	β_3	β_4	β_5
0.2205	0.9765	-0.0439	-0.0134	-0.1333	-0.3776

for cultural influence, the result went as follows:

Table 5: Cultural Influence Regression Coefficient Table.

β_0	β_1	β_2	β_3	β_4	β_5
-0.1960	0.0297	0.7639	-0.4547	0.9051	0.4067

for building reuse value, the result went as follows:

Table 6: Building Reuse Value Regression Coefficient Table.

β_0	β_1	β_2	β_3	β_4	β_5
0.3212	0.4851	0.1841	0.0755	-0.0802	-0.6585

6.4 Goodness-of-fit Test

To evaluate the performance of our multiple linear regression model, we have decided to use model goodness-of-fit measures. Goodness-of-fit measures helped us determine how well the model fits the observed data.

There were several measures of model goodness-of-fit that we can use, including the coefficient of determination (R-squared), adjusted R-squared, and root mean squared error (RMSE). In this case, we have already calculated the R-squared value, which measures the

proportion of variance in the response variable that was explained by the explanatory variables.

To calculate the adjusted R-squared value, we used the following formula:

$$R^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \overline{Y})^2}$$

where n is the sample size, and p is the number of explanatory variables in the model. The adjusted R-squared value is a modification of the R-squared value that adjusts for the number of variables in the model. It penalizes models with more variables, which helps prevent overfitting.

In summary, we have chosen to evaluate our multiple linear regression model using model goodness-of-fit measures. These measures helped us determine how well the model fits the observed data.

Table 8: Table of Goodness-of-fit R^2.

Regression equation	R^2
Y_1	0.9347
Y_2	0.8944
Y_3	0.8896

7 Sensitivity Analysis

Sensitivity analysis is a valuable tool for assessing the robustness and reliability of decision-making models, such as the impact prediction model of Olympic options.

In this chapter, we would discuss the principles and methods of sensitivity analysis, as well as its applications and limitations in the context of the impact prediction model of Olympic options. We would also present several case studies and examples to illustrate how sensitivity analysis can be used to inform decision-making and improve the accuracy and reliability of the model's results.

7.1 Sensitivity

In order to conduct sensitivity analysis, we kept the other independent variables as the average value, change the venue investment by 15%, floor area, number of sponsors, security investment, manpower input, and draw them on a graph. This analysis allowed decision-makers to explore how changes in input parameters or assumptions affect the output of the model, and to identify which factors have the greatest impact on the results. In the context of the Olympics model, sensitivity analysis could help decision-makers to better understand the potential risks and uncertainties associated with hosting the Games, and to evaluate the sensitivity of the model to different scenarios and inputs.

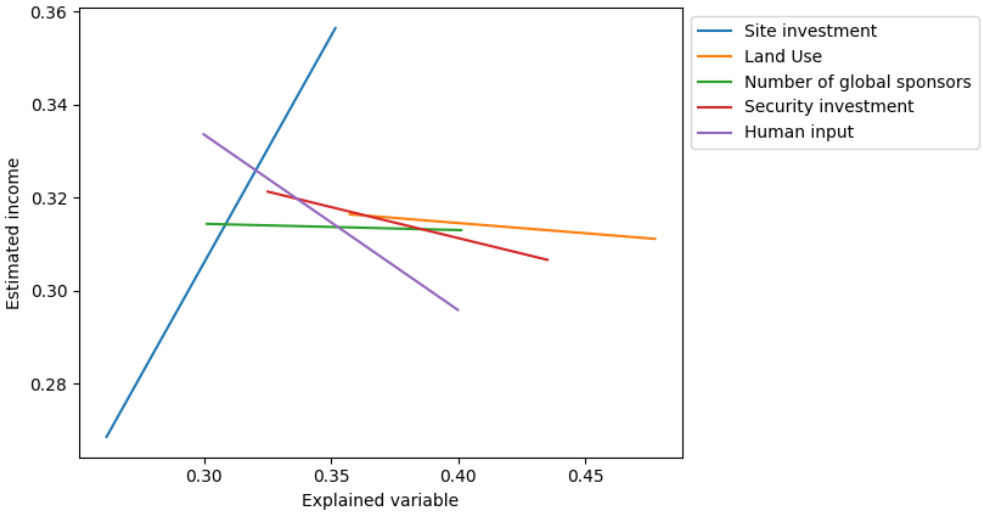

Figure 3. Sensitivity Analysis Regarding Estimated Income.

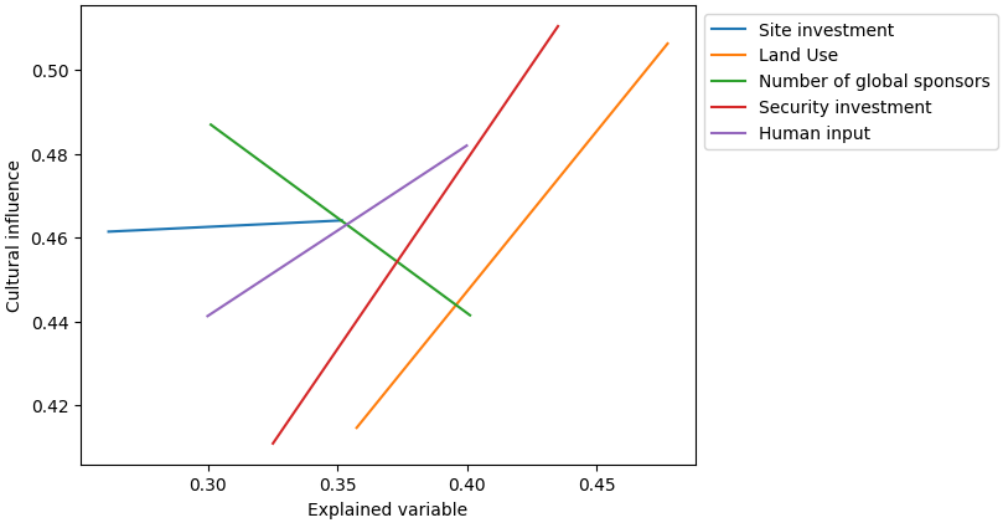

Figure 4. Sensitivity Analysis Regarding Cultural Influence.

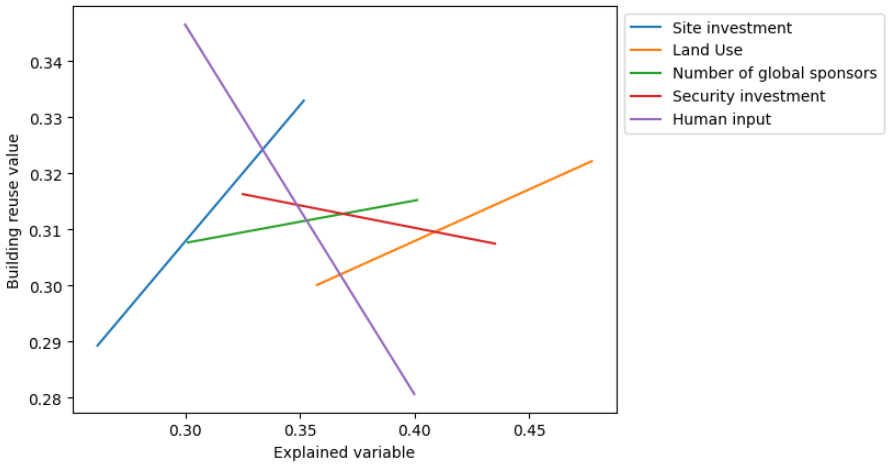

Figure 5. Sensitivity Analysis Regarding Building Reuse Value.

Based on the analysis of Figure 3, Figure 4 and Figure 5, it could be inferred that the site investment exerted the most significant influence on the response variable, i.e., the estimated income. In contrast, the human input had a negative effect on the explained variable. As for the remaining three explanatory variables, namely, land use, the number of global sponsors, and security investment, they had relatively limited impacts on the estimated income.

Regarding the cultural influence model, it could be concluded that security investment and land use had a substantial effect on the predicted cultural influence, followed by the amount of human input. Site investment had a minimal impact on the total output. Furthermore, an increase in the number of global sponsors might lead to a decrease in cultural influence, potentially attributed to a dislike of excessive advertising.

When considering building investment, the importance of site investment remained prominent, with other parameters having either positive or negative impacts.

7.2 Partial Derivative

In order to quantify the impact of a change in input variables on the output, partial derivatives are often employed. The partial derivative is a mathematical tool used to measure the rate of change of a function with respect to a specific variable while holding other variables constant. Specifically, if a function $f(x_1, x_2, \ldots\ldots, x_n)$ is defined, the partial derivative of f with respect to variable x_i is denoted by $\frac{\partial \mathrm{f}}{\partial \mathrm{x}_i}$ and measures how much f changes with respect to x_i, while holding all other variables constant. By using partial derivatives, we can calculate the sensitivity of a model output with respect to each input variable and determine which variables have the most significant impact on the output.

Since for the multiple linear regression model, partial derivatives of all variables are the Corresponding coefficients. So, from the **Table 4**, **Table 5**, and **Table 6** in **section 6.3**, we could obtain these coefficients as our partial derivatives.

(1) For **projected income**, it could be seen that **Land Use** had the greatest impact on the regression equation, while **Number of Global Sponsors** had the smallest impact on the regression equation.
(2) For **cultural influence**, it could be seen that **Security Investment** had the greatest impact on the regression equation, while **Land Use** had the smallest impact on the regression equation.
(3) For **building reuse value**, it could be seen that **Land Use** had the greatest impact on the regression equation, while **Security Investment** had the smallest impact on the regression equation.

8 Model Evaluation and Further Discussion

Model evaluation is crucial for identifying the strengths and weaknesses of the model,

assessing its accuracy and predictive power, and determining whether it is fit for purpose. In the context of the Olympics model, evaluation is especially important, given the high stakes involved in hosting the Games and the potential for significant economic and social impacts.

In this chapter, we will present a comprehensive evaluation of the impact prediction model of Olympic options, including an assessment of its assumptions, limitations. We will also discuss the implications of our findings for decision-making and future research. By evaluating the model in a rigorous and transparent manner, we aim to provide decision-makers with the necessary tools and insights to make informed and evidence-based decisions about hosting the Olympics, and to guide further development and improvement of the model in the future.

8.1 Strengths

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) is a multi-criteria decision-making method that is commonly used in a variety of fields, including engineering, management, and finance. It is a valuable tool for evaluating the performance of different alternatives based on multiple criteria. Here are the Advantages of our TOPSIS with entropy weight:

- TOPSIS with entropy weight can effectively handle decision-making problems that involve multiple criteria or attributes. It allows decision-makers to compare alternatives based on several factors, and provides a rank order of the alternatives.
- Entropy weight helps to address the problem of criteria weighting, which is often a subjective and challenging task. Entropy weight assigns weights to criteria based on their relative importance, and considers both the dispersion and the diversity of the criteria.
- TOPSIS with entropy weight is a transparent and easy-to-use method. It provides a clear decision-making process and can be easily understood by decision-makers.

Multiple linear regression is a statistical technique that allows us to model the relationship between a dependent variable and multiple independent variables. There are several advantages and disadvantages to using multiple linear regression models in our dataset analysis. Advantages of Multiple Linear Regression Model:

- Multiple linear regression models can help us to identify the relative importance of each independent variable in explaining the variation in the dependent variable. This can provide insight into which factors have the most significant impact on the outcome we are interested in.
- Multiple linear regression models can help us to test the significance of the relationship between the independent and dependent variables. This can help us to determine whether the relationship we observe in the data is statistically significant or due to chance.
- Multiple linear regression models can allow us to make predictions about the value of the dependent variable for different values of the independent variables. This can be useful for forecasting and decision-making purposes.

8.2 Weaknesses

- The entropy method mainly uses the variation degree of index data to calculate the weight, and there may still be some deviation from the actual weight.
- Only when the value of the index is within a reasonable range can the linear regression model have a relatively accurate forecasting function. If there are outliers in the dependent variable, the predicted value may be inaccurate.

9 Conclusion

We propose a model to measure the impact of the Olympics and a model to predict the impact the Olympics might bring. At the same time, we demonstrate the feasibility of different strategies and policies from various angles, and discuss the preparation time line and important time points of the Olympic Games and their significance.

References

Authorization of running of special olympics torch relay through Capitol Grounds: Report (toaccompany H. con. res. 81) *(including cost estimate of the Congressional Budget Office)*. (1993). Washington, D.C.?: U.S. G.P.O.

Butterfield, M. (2012). *The Olympics*. London: Franklin Watts.

Dechow, P. M., Lawrence, A., & Luo, M. (2016). Do investors use the Olympics as a category for investment and should they? *SSRN Electronic Journal*. doi:10.2139/ssrn.2803396

Flyvbjerg, B., & Stewart, A. (2012). Olympic proportions: Cost and cost overrun at the Olympics 1960-2012. *SSRN Electronic Journal*. doi:10.2139/ssrn.2238053

Flyvbjerg, B., & Stewart, A. (2016). The Oxford Olympics study 2016: Cost and cost overrun at the games. *SSRN Electronic Journal*. doi:10.2139/ssrn.2804554

Framing: Interpreting the Olympics project. (2012). *Host Cities and the Olympics,* 59-77. doi:10.4324/9780203121337-9

Horne, J., & Whannel, G. (2020). Los Angeles and the Olympics. *Understanding the Olympics,* 54-71. doi:10.4324/9780429345272-4

The IOC, political economy and the games bidding process. (2016). *Understanding the Olympics*, 58-79. doi:10.4324/9781315712529-3

Nelson, M. (2010). Labour geography and the 2010 Vancouver olympics. *Studies in Political Economy, 86*(1), 131-165. doi:10.1080/19187033.2010.11675029

Pawlak, K. (2022). 'volunteer hub' improves volunteer communication, engagement. *The Volunteer Management Report, 27*(5), 3-3. doi:10.1002/vmr.31866

Rimmer, A. (2020). Covid-19: Government appoints former Olympics chief to lead PPE Production drive. *BMJ,* M1580. doi:10.1136/bmj.m1580

Science of the winter olympics: Olympics motion. (2011). *SciVee*. doi:10.4016/37019.01

Sidder, A. (2018). Cycles of mountain building formed 2018 Winter Olympics terrain. *Eos, 99*. doi:10.1029/2018eo092751

Appendices

Appendix 1

Introduce: Data set of Olympic Games over the years

Time	Capital	Country	Site Investment	Olympic revenue	GDP	GDP growth in 8 years
1992	Barcelona	Spain	96.9	97.3	0.631	3.6
1992	Albertville	France	NA	17	1.401	NA
1994	Lillehammer	Norway	22.3	24	0.127	-0.6
1996	Atlanta	U.S.	41.4	51	8.073	2.7
1998	Nagano	Japan	22.3	14	4.098	1.4
2000	sydney	Australia	50.3	53.9	0.416	4.2
2002	Salt Lake City	U.S.	25.2	25	10.93	4
2004	athens	Greece	23.4	23	0.241	4.1
2006	Turin	Italy	43.7	32	1.95	1
2008	Beijing	China	68.1	30.14	4.594	10.2
2010	Vancouver	Canada	25.4	51	1.617	3.8
2012	London	U.K.	149.6	119	2.706	0.8
2014	Sochi	Russia	218.9	160	2.059	-3.7
2016	Rio de janeiro	Brazil	45.8	31	1.796	-3.5
2018	Pyeongchang	South Korea	130	76	1.725	
2020	Toyko	Japna	250	70	4.941	NA
2022	Beijing	China	39	31.1	18.321	NA

Relative benefit	Land Use	Viewers	Travel Growth Rate (Host City)	Venues utilization rate	Media coverage (number of Google search entries)
0.633914422	538	35.4	NA	4	38.3
NA	NA	12.5	0.951219512	2	55.8
13.38582677	200	17.5	1.088888889	1	10.2
1.189149015	267	33.6	1.101449275	2	37.6
-2.025378233	NA	18.5	NA	3	32.2
8.653846154	227	39.3	1.076923077	4	55.5
-0.018298262	1,063	18.8	1.285714286	2	42.3
-1.659751037	2,400	40.7	1.305555556	3	41.8
-6	2,000	21.3	NA	2	45.4
-8.262951676	1,177	45.6	0.978236142	5	53.5
15.83178726	1,000	21.4	1.283055199	3	57.4
-11.30820399	2,500	38.8	1.025	5	62.8
-28.60611948	NA	17.6	1.095485547	2	62.4
-8.240534521	3,000	36.5	1.04437401	2	101
-31.30434783	NA	24.7	1.311844078	2	91.9
-36.4298725	5,200	NA	0.931873479	NA	121
-0.43119917	NA	32.3	NA	NA	158

MEMORANDUM OF UNDERSTANDING

Date	April 1st, 2023
To	International Olympic Committee (IOC)
From	Group 2333325
Subject	Recommendations for Hosting a Successful Olympic Games

Introduction

The purpose of this memorandum is to provide recommendations to the International Olympic Committee (IOC) regarding the hosting of a successful Olympic Games. Our proposed algorithm has utilized the **TOPSIS with entropy weight method** to establish indicators for the impact of hosting the Olympic Games from various perspectives, including economic, land use and etc. We have also considered feasibility, implementation timeline, and potential strategies that could impact these indicators using the multiple linear regression. We first filter out some strategies in the "solution space" by strictly dominating the strategy. Then we need to establish three **linear regression** equations to find the relationship between input and output, which requires looking for data on the Internet.

Strategy:

we started by selecting a set of solutions from the "solution space" using strict dominance strategy. Then, we delved into the details of our solutions by using mathematical modelling to relate four independent variables: venue investment, land area, number of global sponsors and advertisement investment (which were identified as important indicators using the TOPSIS method) with three dependent variables: expected revenue, cultural influence, and building reuse value. We assessed the significance of relationships to determine the most effective solution.

Throughout the process, I applied rigorous statistical techniques and my expertise in modeling to ensure the validity and reliability of our findings.

Time Frame:

According to our estimation, the following planning would be the most suitable one.

Time Frame 1	Security preparations
Time Frame 2	Security preparations
Time Frame 3	Completion of relevant law
Time Frame 4	Start promoting the Olympic Games
Time Frame 5	Completion of basic construction of all venue
Time Frame 6	Testing of the Venues; Preselling the tickets

Policy Recommendation

1. **Economic Impact**: The economic impact of hosting the Olympic Games can be significant, but it is important to carefully consider the costs and benefits of hosting the Games to ensure that they are financially sustainable in the long term.
2. **Land Use**: Hosting the Olympic Games requires a significant amount of land, which can have both positive and negative impacts on the host city. It is important to consider the long-term impact of land use decisions, including environmental and social impacts.