Problem Chosen	2023 MCM/ICM Summary Sheet	Team Control Number
Z		2333779

Beyond the Status Quo: Exploring Innovative Ways to Host the Olympic Games

Summary

Nowadays fewer cities may be interested in hosting the Olympic Games due to various reasons like huge cost of venues, further maintenance fee or some possible negative effects on the environment. The Olympic Games are a significant event that brings together athletes, fans, and countries from around the world. The spirit of the Olympics is more than just winning medals. It's about celebrating the values of excellence, friendship, and respect, and bringing people together from around the world to share their joint joy in sports. However, hosting the Olympics can be a challenging and costly endeavor. Therefore, new explorations and breakthroughs in the existing mode of the Olympic Games are necessary.

This article proposes a model to assess **the potential impacts** that hosting the Olympics may have on a city's overall situation. The article aims to offer recommendations to alleviate the burden of hosting the Olympics and ensure that it can function as a reminder of peace in the future as it used to be.

To determine **the quantitative impact** of hosting the Olympics, the article adopts the **Polynomial Curve Fitting Method** based on relevant data that may be impacted by the Olympics. The chosen **economic aspects** are **tourism revenues** and **construction fees**, while **GDP per unit of land** and **greenhouse-gas emissions** represent **the host city's potential** and **sustainability**, respectively. **The numbers of appearances** of the Olympic city in the news websites represent the aspect of **host city prestige**. Using the collected data, we construct an **Ideal City Model (IC-Model)** to measure the impacts of hosting the Olympics on the city.

We use the **Entropy Weight Method** to calculate and compare the **variability** of the weights of each aspect in the **IC-Model**. The economic aspect takes **the largest share**, followed by the potential and prestige aspects. We then define S_{tk} as the total cost of hosting the Olympics and import six groups of data into the **IC-Model**. The analysis of S_{tk} provides insight into the city and event regimes that may have the lowest cost.

In conclusion, the article presents a model that assesses the potential impacts of hosting the Olympics on a city's overall situation. We recommend that the IOC consider the **economic, city-potential related, and city-prestige related aspects** of the host city when selecting a venue for the Olympics. We suggest that the IOC focus on cities that have the potential to sustain fewer loss, as measured by the **IC-Model**. Finally, We offer recommendations to the ICMG based on our analysis. Overall, this article provides a framework for **assessing the impact** of hosting the Olympics and recommends strategies to alleviate the burden and maximize the benefits of hosting the event.

Keywords: Polynomial Curve Fitting Method Entropy Weight Method Scatter Chart

Contents

1 Introduction

1.1 Problem Background

The Olympic Games are a significant event that brings together athletes, fans, and countries from around the world. However, hosting the Olympics can be a challenging and costly endeavor. The International Olympic Committee (IOC) is facing a decreasing number of bids to host the Olympics – both Summer and Winter Games, as people start to cast doubt on its significance and economical and social benefits. There are nine candidate cities that bid for the 2012 Summer Olympic Games, seven for the 2016 Summer Olympics, five for the 2020 Summer Olympics, and four for the 2024 Summer Olympics[1]. On September 13, 2017, the International Olympic Committee directly awarded the 2028 Summer Olympics to Los Angeles without even calling for other competitors. As a consequence, some new strategies and options are necessary to ensure that the Olympics remain influential and spread its idea of peace.

1.2 Restatement of the Problem

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

- Task 1: Identify the diverse factors of a city that are affected by hosting the Olympic Games. This requires conducting extensive research on previous Olympic Games to determine the factors that had the most significant impact on the city. Once the factors have been identified, the next step is to select the most influential ones and explore their mathematical impact. This involves using statistical analysis techniques to quantify the impact of each factor and determine their relative importance.

- Task 2: Construct a model that can assess the extent to which the Olympic Games alter a city concerning the dominant factors. To accomplish this, we will employ the **Entropy Weight Method**, which is a mathematical technique used to measure the impact of each factor. The model will be constructed by collecting data on the various factors that are affected by hosting the Olympic Games and using the **Entropy-weight Method** to determine their relative importance.

- Task 3: Investigate how the total cost of hosting the Olympic Games can be minimized by modifying current organizing policies. This task involves examining the effects of various policy changes, such as reducing the number of events or building temporary facilities instead of permanent ones. By analyzing the cost-effectiveness of each policy change, we can provide useful recommendations to the International Olympic Committee (IOC) on how to minimize the cost of hosting the Olympic Games.

1.3 Our Work

We represent our work by a picture as follows1

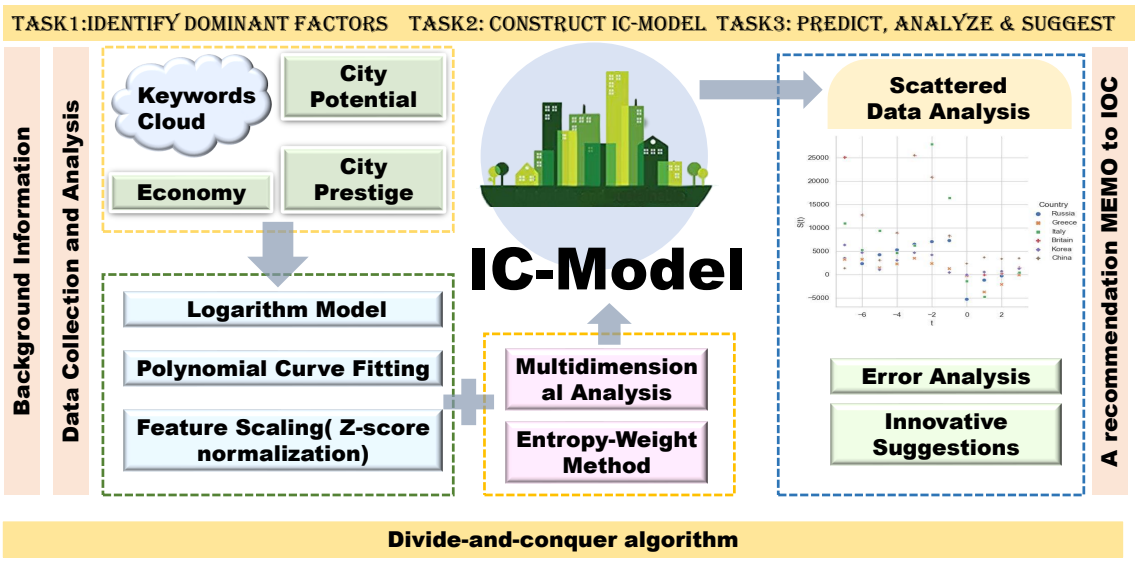

Figure 1: Our Work

2 Assumptions and Justifications

We make the following fundamental assumptions that are each correctly justified in order to simplify the issue:

- **Assumption 1** We suppose the change of every relevant factor in the model is totally due to the Olympic Games and other possible reasons are omitted in order to keep the model focused enough on the effect of the Olympic Games.
  - **Justification 1:** Because at present, holding Olympic Games is a rare and very significant event for a city, which means on most occasions, the city will not have much more energy for other mega-events providing a similarly significant effect to Olympic Games during this period on the city. Besides, for a considerable period of time after and before the Olympic Games, these factors concerning the development of a city are basically long-term and stable. Therefore, we can just build up a city model developing at the pace with historical experience and set the holding of the Olympic Games as the only independent variable.
- **Assumption 2** We suppose there are no dramatic changes in the following years of the Olympic Games, such as wars, economic crises, natural disasters, etc.
  - **Justification 2:** To study the impacts on different aspects, such as the economics, the society, and the human satisfaction of the city, we should focus on the corresponding results of the Olympics, which are generally controllable and mild, but these dramatic affairs all cause immense changes and they will invalidate our model.

- **Assumption 3** We suppose that the Olympic Games will remain the biggest sports activity in the world and its own popularity will not shrink significantly among the pop even though governments tend to not host them.
  - **Justification 3:** If we want to study the impact of the Olympic Games in the long run, we should make sure that even if the IOC set a permanent location for the games, it will preserve its influence and attract enough people, though the total number be may smaller than before, and so our model and recommendations are meaningful.

3 Notations

Symbol	Definition
t	the interval between the target year and the Olympic year (t was considered 0 on the Olympic year)
T_{tk}	Tourism revenues of the city k in year t
G_k	the Gross Domestic Product of the city k
C_k	The construction fee of the Olympic revenue of the city k
Pr_{tk}	The prestige of the city k in the year t
Po_{tk}	The city potential of the city k in the year t
Ge_{tk}	The greenhouse emission of the city k in the year t
L_{tk}	The land use of the city k in the year t
S_{tk}	The overall indicator of the city k in the year t
E_{tk}	The economy indicator of the city k in the year t
N_{tk}	The appearances of the news related to the city k in the year t

4 Data Acquisition and Processing

4.1 Data Acquisition

We use Google Scholar to find the frequency of occurrence of keywords pertaining to the city's factors influencing the Olympic Games.

We look into precisely what the dominant factors of our chosen indicators is and how they affect it. Then, on the basis of our plan of which data we need to use, we go to the website *https://insights.ceicdata.com/Untitled-insight/views* to acquire information about the change in tourism-related GDP in a city and the greenhouse gas emission due to the holding of the Olympic Games when we look further into precisely what affects the chosen dominant factors have on the city hosting the Games and how it occurs.

4.2 Data Cleaning

We try to ensure the feasibility of the data we have collected and do the process of data cleaning, including dropping the repetitive data , deleting or filling the rows that the data is missing select and

categorizing the data we really need to use by groups. For example, when we are processing the data of GDP, we fill in some of the blanks with the average of its nearest two elements. Additionally, when we get the GDP data of the Olympic host cities, we abandoned all the data after 2019, because the COVID-19 impact becomes inevitable and so that these data doesn't in parallel with our assumptions. After we clean our data, we can get receive better fitting results and can be more approached the reality.

4.3 Data Processing

We create a cloud chart based on this frequency with varying font sizes to represent the significance of various factors, assisting us in identifying the predominate factors. 2:

Figure 2: Key Words Cloud

Based on Key Words Cloud, we can easily get the dominant factors participating in the process of affecting the city. However, the keywords related to Olympics and city-effects have not been classified and their coverage may overlap somewhat, which is not helpful for further concrete analysis of how they act on cities in a quantitative way. Therefore, we divide these keywords into two levels and categorize secondary factors into primary factors or summarize the factors reflecting similar effects together to streamline the number of independent variables we will use to rationalize our model. The primary factors are economy, city potential and city prestige and our process of analysis of these factors is shown in the figure Analysis of dominant participating factors. 3:

In reality, it is sure that the Olympics have a long-term impact on the host city's economy. But there are precedent researches which demonstrated that the dominant economy benefit is from the tourism, caused by spectators who mainly stimulate the city tourism in the two-weeks period of the game. Another kind of tourists is called "induced tourism", which is generated by heightened international

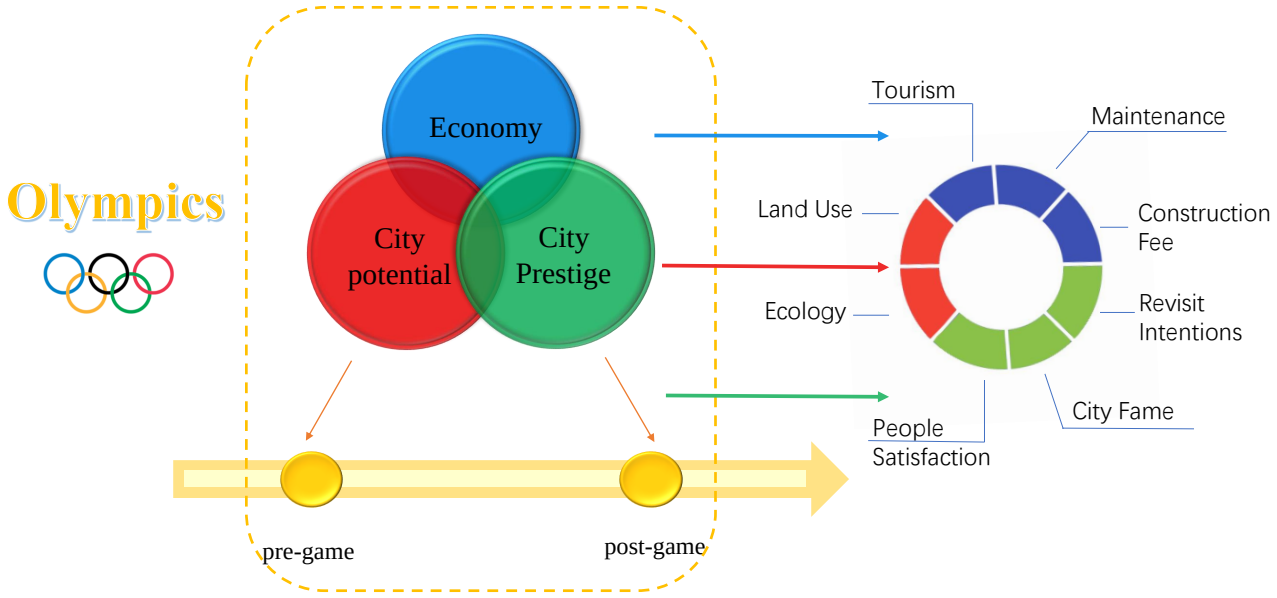

Figure 3: Analysis of dominant participating factors

awareness of the host city and the host country through international media coverage of the Games, and pre-Games stories.[9]

The dominate expenditures caused by the Olympic Games are the operation of the Games over the three years of preparation prior to the Games, and in the two weeks of Olympic events construction of the Games site. Another expenditure includes the construction of the sporting venues, accommodation for participants(Olympic Village) , and necessary upgrading of transport infrastructure to make sure that the traffic system is able to bear the large volume of visitors. All those impacts are in the period of 12 years that begins from the announcement of the hosting, when there is 7 years before the Games and ends after 4 years, when the next Olympic Games begin[2]. So the duration we investigate on is from t=-7 to t= 4.

After the deep analysis of the three major influencing aspects, we found that we could use similar ways to solve its relationship with the information of city itself and the time t. We thus use **Divide-and-conquer algorithm** to divide the building of our whole models into three aspects and solve them step by step using the similar ways. After solving the smaller questions, we combine them and build our final model.

5 Task1 Solution: Identify Dominant Factors

In order to explore the quantitative impact of hosting Olympics Games on the city, we introduce **Polynomial Curve Fitting**2, based on the relevant data from some cities which have held Olympic Games, to achieve the mathematical relationship between time and the indicators used to reflect primary factors.

A polynomial fit is a linear model in which the relationship between the model and the fitted parameters is linear. The goal of polynomial fitting is to construct a polynomial function of order of the input such that the polynomial approximates the relationship between the input and output. Although the relationship between the independent variable and the dependent variable is not necessarily polynomial, it is always possible to approximate the relationship between input and output using a sufficient number of orders. Compared to the linear fitting, which has only two parameters, the polynomial curve fitting has less error.

Since polynomial curve fitting is about constructing a polynomial function, the higher degree of the independent variables will lead to more accuracy but when the degree is too high, there is possibility for over-fitting. To evaluate the extent of fitting accuracy, we introduce the coefficient of determination R^2 to reflect what percentage of the fluctuations in the dependent variable can be described by the fluctuations in the independent variable and when R^2 is closer to 1, the fitting accuracy is higher.

When dealing with the raw data, we take the logarithm of them based on **logarithm model**, which is usually applied to systematically evaluate the relationships and magnitude of interactions among variables between purely defined classes of variables.

Besides, we use **feature scaling(Z-score normalization)**[1] to make the values of independent variables into the range [-1,1] to reduce difference of the orders of magnitude, which is conducive to an accurate fit.

$$(X)' = \frac{X - \bar{X}}{\sigma} \tag{1}$$

5.1 Economy

The indicators used to reflect the economic impact of hosting the Olympics are tourism revenues and the construction fees for the Olympic site. Research has shown that the host city experiences a surge in tourism during the event, making tourism revenue an important indicator of economic impact[6]. Additionally, the construction fees for the Olympic venues are typically large compared to the later maintenance costs, making them a significant factor to consider. Therefore, by using tourism revenues and construction fees, we can comprehensively evaluate the economic impact of the Olympic Games on the host city.[3].

In order to ensure a suitable degree for the objective function, we create a for-loop with python from degree 1 to degree 100 to test and we finally find that when the degree is 2,R^2 about the relationship between tourism revenues and time is closest to 1 and when the degree is 2, R^2 about the relationship between construction fee and time is closest to 1.

We construct an equation to reflect the tourism revenues of the city *k* at *t* point.

$$T_{tk} = f(T_{(t-1)k}, t)$$

By adding the degree, we can get the polynomial fit equation:

$$\begin{aligned} T_{t_k} = intercept + coefficient[0] \times {T_{(t-1)_k}}^0 t^0 + coefficient[1] \times {T_{(t-1)_k}}^1 t^0 \\ + coefficient[2] \times {T_{(t-1)k}}^0 t^1 + coefficient[3] \times {T_{(t-1)k}}^2 t^0 \\ + coefficient[4] \times {T_{(t-1)k}}^1 t^1 + coefficient[5] \times {T_{(t-1)k}}^0 t^2 \end{aligned} \tag{2}$$

By importing the data of tourism revenues of different countries, including Korea, Italy, Britain, Brazil, Canada and China in the period of the Olympic Games, we can get the coefficient matrix of the Equation(2) with help of the library function **sklearn** in python.

The data of tourism revenues of different Olympic cities is4:

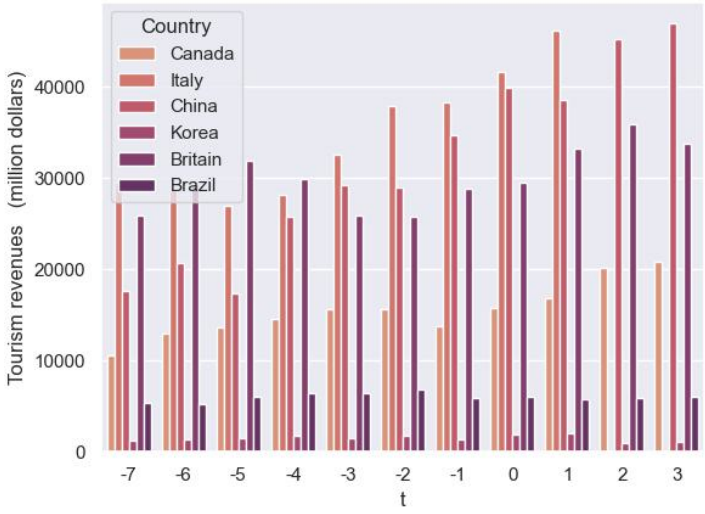

Figure 4: Tourism revenues of different Olympic countries

The R^2 of every Olympic city that we analyze are shown in the table: Evaluation of polynomial curve fitting of tourism revenues 1

Country with Olympic city	R^2
$Korea$	0.7370
$Italy$	0.7812
$Britain$	0.6737
$Brazil$	0.6276
$China$	0.3485
$Canada$	0.6307

Table 1: Evaluation of polynomial curve fitting of tourism revenues

According to Table1, the R^2 of each Olympic city is near 1, which shows that the polynomial curve fitting of tourism revenues and time is successful. By taking the average of these coefficient matrices we can identify the unknown numbers of the function related to the factor of tourism revenues in the model we have constructed.

Coefficient matrix:

$$\begin{pmatrix} 0.0000 & 0.5170 \\ -0.0503 & 0.3769 \\ -0.0667 & -0.5653 \end{pmatrix}$$

Intercept: -0.0354

Complete equation for polynomial curve fitting of tourism revenues

$$\begin{aligned} T_{tk} = 0.0354 + 0.5170 \times {T_{(t-1)k}}^{1} t^{0} + 0.0503 \times {T_{(t-1)k}}^{0} t^{1} \\ +0.3769 \times {T_{(t-1)k}}^{2} t^{0} - 0.0667 \times {T_{(t-1)k}}^{1} t^{1} - 0.5653 \times {T_{(t-1)k}}^{0} t^{2} \end{aligned} \tag{3}$$

Since construction fee is only affected by the holding of Olympics, which is an instantaneous variable, we don't include the variable t when constructing the equation of the construction fee of the city k.

$$C_k = f(G_k)$$

By adding the degree we can get the polynomial fit equation:

$$\begin{aligned} C_k = intercept + coefficient[0] \times {G_k}^{0} + coefficient[1] \times {G_k}^{1} \\ +coefficient[2] \times {G_k}^{2} \end{aligned} \tag{4}$$

By importing the data of construction fee of different Olympic cities in different countries, including Korea, Italy, Britain, Brazil, Canada and China in the period of the Olympic Games, we can get the coefficient matrix of the Equation(4) with help of the library function **sklearn** in python.

The data of construction fee of different Olympic cities is5:

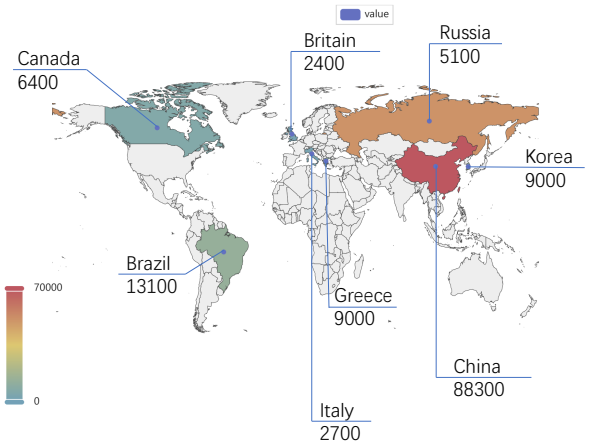

Figure 5: Construction fee(million dollars) of different Olympic cities

The R^2 of the analysis of construction fee of each Olympic city is 0.7025, which shows that the polynomial curve fitting of tourism revenues and time is successful. Therefore, we can identify the unknown numbers of the function related to the factor of construction fee in the model we have constructed.

Coefficient matrix:

$$\begin{pmatrix} 0.0000 & 0.5040 & 0.3022 \end{pmatrix}$$

Intercept: -0.2644

Complete equation for polynomial curve fitting of construction fee

$$C_k = -0.2644 + 0.5040 \times {G_k}^1 + 0.3022 \times {G_k}^2 \tag{5}$$

5.2 City Potential

Upon analyzing the secondary factors, it can be concluded that land use and ecology can be combined to assess the potential of a city's development. Hosting the Olympic Games has a significant impact on land use due to the construction of new sports facilities, transportation infrastructure, and other related developments. The ecological impact of hosting the Olympic Games is also crucial, as it can result in substantial environmental impacts such as carbon emissions and alterations to green spaces.[4]

Hosting The Olympic Games results in a significant change in land use in the host city both before and after the event. This includes the construction of new sports facilities, transportation infrastructure, and other related developments. To comprehensively assess the impact of the Olympics on the land use of a city, we chose to use GDP per unit of land area. This metric reflects the revenue-related effects of the infrastructure changes, which are likely to have a long-term impact on the land.

We construct an equation to reflect the GDP per unit of land area of the city *k* at *t* point.

$$L_{tk} = f(t)$$

By adding the degree=3 we can get the polynomial fit equation:

$$L_{tk} = intercept + coefficient[0] \times t + coefficient[1] \times t^1 + coefficient[2] \times t^2 \tag{6}$$

By importing the data of the GDP per unit of land area of different Olympic cities in different countries, including Canada, Brazil, Russia, Korea, Italy and China in the period of the Olympic Games, we can get the coefficient matrix of the Equation(6) with help of the library function **sklearn** in python.

The data of the GDP per unit of land area of different Olympic cities is6:

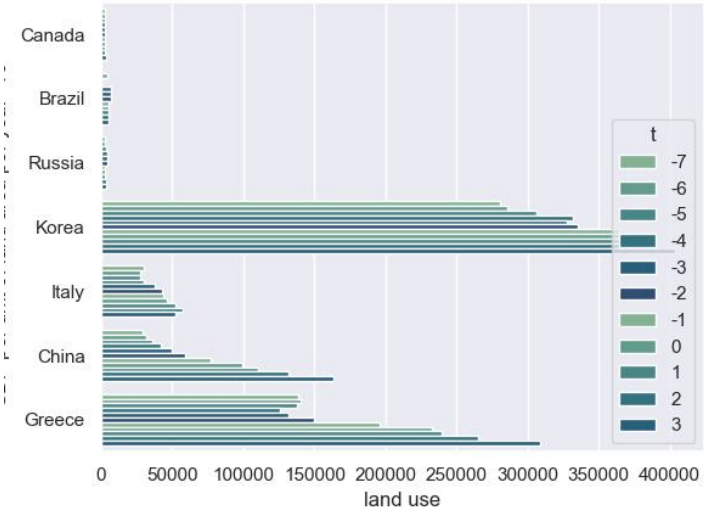

Figure 6: The GDP per unit of land area (million dollars /square kilometer)of different Olympic cities

The R^2 of every Olympic city that we analyzes are shown in the table: Evaluation of polynomial curve fitting of the GDP per unit of land area 2

Country with Olympic city	R^2
$Canada$	0.8353
$Brazil$	0.7513
$Russia$	0.3284
$Korea$	0.5713
$Italy$	0.4962
$China$	0.6711

Table 2: Evaluation of polynomial curve fitting of the GDP per unit of land area

According to Table2, the R^2 of each Olympic city is near 1, which shows that the polynomial curve fitting of the GDP per unit of land area is successful. By taking the average of these coefficient matrices we can identify the unknown numbers of the function related to the factor of greenhouse emissions in the model we have constructed.

Coefficient matrix:

$$\begin{pmatrix} 0.0000 & 2.3828 \\ 0.8686 & -0.8175 \end{pmatrix}$$

Intercept: -0.1874

Complete equation for polynomial curve fitting of the GDP per unit of land area

$$L_{tk} = -0.1874 + 2.3828 \times t + 0.8686 \times t^1 - 0.8175 \times t^2 \tag{7}$$

The hosting of the Olympics typically involves the construction of new infrastructure, transportation systems, and housing for athletes and spectators, which can result in significant greenhouse emissions. The construction phase alone requires large amounts of carbon-intensive materials, such as concrete and steel, and the energy needed to transport and process them. Furthermore, the influx of tourists and athletes during the event can exacerbate local transportation systems, leading to increased traffic congestion and emissions from cars and buses. The energy used to power the Olympic venues, including lighting, heating, and cooling, can also contribute significantly to the carbon footprint of the event.

We construct an equation to reflect the greenhouse emissions of the city *k* at *t* point.

$$Ge_{tk} = f(Ge_{(t-1)k}, t)$$

By adding the degree=2 we can get the polynomial fit equation:

$$\begin{aligned} Ge_{tk} = intercept + coefficient[0] \times {Ge_{(t-1)k}}^0 t^0 + coefficient[1] \times {Ge_{(t-1)k}}^1 t^0 \\ + coefficient[2] \times {Ge_{(t-1)k}}^0 t^1 + coefficient[3] \times {Ge_{(t-1)k}}^2 t^0 \\ + coefficient[4] \times {Ge_{(t-1)k}}^1 t^1 + coefficient[5] \times {Ge_{(t-1)k}}^0 t^2 \end{aligned} \tag{8}$$

By importing the data of greenhouse emissions of different Olympic cities in different countries, including Korea, Italy, Britain, Brazil, Canada, China, Russia and Greece in the period of the Olympic

Games, we can get the coefficient matrix of the Equation() with help of the library function **sklearn** in python.

The data of greenhouse emissions of different Olympic cities is7:

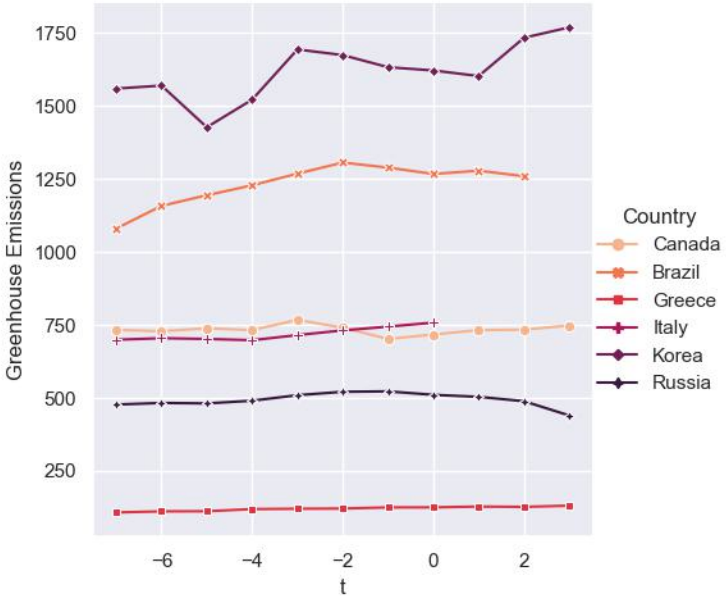

Figure 7: Greenhouse emissions(million tons) of different Olympic cities

The R^2 of every Olympic city that we analyzes are shown in the table: Evaluation of polynomial curve fitting of tourism revenues 3

Country with Olympic city	R^2
$Canada$	0.8608
$China$	0.4930
$Brazil$	0.5886
$Greece$	0.6827
$Italy$	0.5910
$Korea$	0.7751
$Russia$	0.6415
$Britain$	0.5402

Table 3: Evaluation of polynomial curve fitting of greenhouse emissions

According to Table2, the R^2 of each Olympic city is near 1, which shows that the polynomial curve fitting of greenhouse emissions and time is successful. By taking the average of these coefficient matrices we can identify the unknown numbers of the function related to the factor of greenhouse emissions in the model we have constructed.

Coefficient matrix:

$$\begin{pmatrix} 0.0000 & 0.4508 \\ 0.1475 & 0.4870 \\ 0.3056 & 0.4217 \end{pmatrix}$$

Intercept: -0.0188

Complete equation for polynomial curve fitting of greenhouse emissions

$$\begin{aligned} Ge_{tk} = -0.0188 + 0.4508 \times {Ge_{(t-1)k}}^1 + 0.1475 \times t^1 + 0.4870 \times {Ge_{(t-1)k}}^2 \\ +0.3056 \times {Ge_{(t-1)k}}^1 t^1 + 0.4217 \times t^2 \end{aligned} \tag{9}$$

5.3 City Prestige

When analyzing the secondary factors, we find that the factors of people satisfaction, revisit intentions and city fame can be summarized as the city prestige.

It is hard to obtain the relevant quantitative information about the impact of hosting the Olympics on people's satisfaction or revisit intentions because this factor is too subjective[7]. However, based on former research in this aspect, we can conclude that the impact is both positive and negative.

On one hand, the excitement and pride of hosting a major international event can create a sense of unity and pride among the local population. On the other hand, the preparation and execution of the Olympics can be disruptive, with potential negative effects on the daily lives of citizens, such as traffic congestion, displacement of local residents, and increased security measures. The financial burden of hosting the Olympics can also lead to budget deficits and increased taxes, which can have a negative impact on people's overall satisfaction[5].

Therefore, we just focus on the factor of city fame, which can be intuitively reflected by numbers of appearances of the Olympic city in the news website per year and the only independent variable is *t*.

Based on these considerations, we construct an equation to reflect the prestige cost of the city *k* at *t* point.

$$Pr_{tk} = -N_{tk}$$
$$N_{tk} = f(t)$$

By adding the degree=3, we can get the polynomial fit equation:

$$\begin{aligned} N_{tk} = intercept + coefficient[0] \times t^0 + coefficient[1] \times t^1 \\ +coefficient[2] \times t^2 + coefficient[3] \times t^3 \end{aligned} \tag{10}$$

By importing the data of numbers of appearances in the news website per year of different Olympic cities, including Korea, Italy, Brazil, Canada, Russia and Greece in the period of the Olympic Games, we can get the coefficient matrix of the Equation(10) with help of the library function **sklearn** in python.

The data of numbers of appearances in the news website per year of different Olympic cities is8:

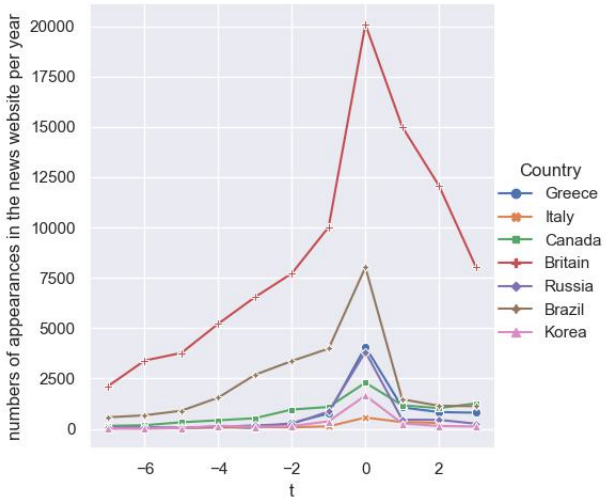

Figure 8: Numbers of appearances in the news website per year of different Olympic cities

The R^2 of every Olympic city that we analyze are shown in the table: Evaluation of polynomial curve fitting of city prestige 4

Country with Olympic city	R^2
$Greece$	0.4348
$Italy$	0.6921
$Canada$	0.7261
$London$	0.8564
$Russia$	0.3699
$Brazil$	0.5423
$Korea$	0.3811

Table 4: Evaluation of polynomial curve fitting of city prestige

According to Table, the R^2 of each Olympic city is near 1, which shows that the polynomial curve fitting of greenhouse emissions and time is successful. By taking the average of these coefficient matrices we can identify the unknown numbers of the function related to the factor of city prestige in the model we have constructed.

Coefficient matrix:

$$\begin{pmatrix} 0.0000 & 1.4025 \\ -0.3652 & -0.5530 \end{pmatrix}$$

Intercept: -0.3166

Complete equation for polynomial curve fitting of city prestige

$$N_{tk} = -0.3166 + 1.4025 \times t - 0.3652 \times t^2 - 0.5530 \times t^3 \tag{11}$$

6 Task 2 Solution: Construct IC-Model

In Task2, we construct our IC-Model by combining the subjective and the objective weights. Our object is to evaluate the total impact of the Olympic games to a city synthetically , to estimate whether and how much the city's comprehensive expenditure to host an Olympic. According to our model and the existing conclusions of the task1, the overall impact can be determined.

6.1 Construct the indicators

In Task1, we have analyzed many influential factors and concluded three decisive ones, economy, city potential, and city prestige. Then, we analyze the indicators from multiple dimensions.

In the **multidimensional analysis**, firstly, in terms of economy factors, T_{tk} and C_{tk} have the same unity (million dollars) and the indicator of the economy cost on the city during the Olympics is the subtraction of the tourism revenues from the construction fee of the Olympics stadiums. Therefore, we define

$$E_{tk} = C_{tk} - T_{tk} \tag{12}$$

as an economic indicator.

Next, when it comes to the indicator of the city potential, we consider the Pr_{tk} as the weighted sum of land use and ecology and decide the weights after feature scaling them.

$$Po_{tk} = w_1 * \frac{L_{tk} - \bar{L_{tk}}}{\sigma_1} + w_2 * \frac{Ge_{tk} - \bar{Ge_{tk}}}{\sigma_2} \tag{13}$$

For the consideration of the weights of the two indicators, we weigh the contribution of the two influencing factors to the city potential. Considering the urban development room as a more important factor than the ecology, we thus decide w_1 = 0.7 and w_2 = 0.3.

In order to construct our final **Ideal City Model(IC-Model)**, we use S_{tk} to measure all the impacts of the Olympics on the city. To identify the relationship between the S_{tk} and the dominant factors (Economy, City Potential and City Prestige), we introduce the weight matrix W_{tk},

Finally, for the indicator of the city prestige, as mentioned in the task 1, we consider it as the negative version of number of the appearances of the city in the news website. Generally, the prestige cost is negative since the prestige of the city always becomes higher during the Olympics.

$$S_{tk} = W_{tk} * \begin{bmatrix} E_{tk} \\ Pr_{tk} \\ Po_{tk} \end{bmatrix} \tag{14}$$

6.2 Determine the weights

In the first part of Task2, we have built matrix for the impacts of hosting the games from various points of views. Now we will utilise the **Entropy Weight Method** to determine the different weight of different aspects that we have researched in task1.

The Entropy Weight Method is a statistic method to determine the weight of indexes objectively by calculating and then comparing the variability of each index. Information is a measure of the degree of order in a system, while entropy is a measure of the degree of disorder in a system. If the information entropy of an index is smaller, the amount of information(and certainty) provided by the indicator is

larger. Thus, this indicator should play an more important role in comprehensive evaluation, and the weight should be higher.

To utilise the Entropy Weight Method, we should wipe out the dimensions of the indexes. As is mentioned above, there are 3 major indexes,

$$E_{tk}, Pr_{tk}, Po_{tk}$$

We choose 5 cities to calculate.

For

$$i \in \{1, 2, 3, 4, 5\}$$

$$X_{ij} = \begin{bmatrix} E_1 & Pr_1 & Po_1 \\ E_2 & Pr_2 & Po_2 \\ E_3 & Pr_3 & Po_3 \\ E_4 & Pr_4 & Po_4 \\ E_5 & Pr_5 & Po_5 \end{bmatrix}$$

Before we standardize the data, since we observe that there is a negative correlation between the prestige and S(t), we need to forwardize the indicator, turning it into an indicator that its increase lead to the increase of the overall indicator.

$$Y_{ij} = \frac{X_{ij} - min(X_i)}{max(X_i) - min(X_i)} \tag{15}$$

Next, we proceed the data standardization after these three indicators are all correlated with the overall indicator positively. To calculate the variability of every index, we can introduce a new P_{ij} , the proportion of every Y_{ij} in the Y_i index.

$$p_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{5} Y_{ij}}, i = 1, 2, 3, 4, 5; j = 1, 2, 3 \tag{16}$$

Then, according to the definition of the information entropy,we can calculate the information entropy E_j by

$$E_j = -\ln(5)^{-1} \sum_{i=1}^{5} p_{ij} \ln p_{ij}, p_{ij} \neq 0 \tag{17}$$

$$E_j = []$$

Finally, the weight of every index w_j is:

$$w_j = \frac{1 - Ej}{3 - \sum E_j} \tag{18}$$

By applying the Entropy Weight Method to our former results from 6.1, we successfully get the weight matrix of the three indicators,

$$w = \begin{bmatrix} 0.7758 & 0.1340 & 0.0902 \end{bmatrix}$$

7 Task3 Solution : Predict, Analyze & Suggest

7.1 Scattered Data Prediction

In task2, we have used the **Entropy Weight Method** to build the IC-Model and to get the overall indicators S_{t_k}, a function related only to the year *t* and the city *k*. The sum of S_{tk} showed the total cost of Olympics, including every indicator factors in the S_{tk} matrix, for the specific Olympic city *k*.

Therefore, by substituting different values of the independent variables *t* and *k* into the S_{tk} matrix we can predict how the changes of current Olympic policies will alter the effects on Olympic Games on a country.

We import the data of S_{tk} of six countries that have held Olympic Games, including Russia, Greece, Italy, Britain, Korea and China and show them in a scatter chart9:

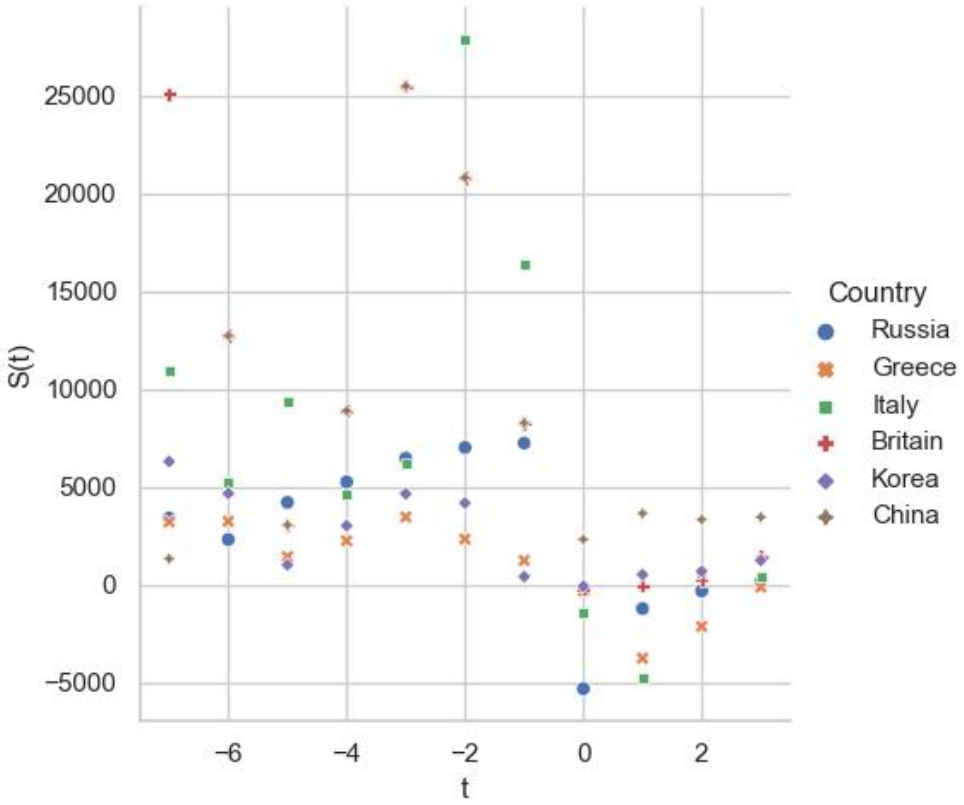

Figure 9: S-t Scatter Chart

When the sum of S_{tk} is getting close to 0 from positive numbers, the cost of the Olympics on the city is less because every S_{tk} reflects the cost of it in the particular year.

According to Figure9, we can find that generally when *t* is near 0 point, about (0,3), the total cost of hosting Olympic Games is the least. Possible explanation for this is that when *t* is near 0, it is the time during the Olympic Games, when tourism is bringing maximum dividends4. Besides, this is the time with the lowest extra cost like maintenance fee. Therefore, it's reasonable to conclude that when *t* is near 0 point the total effect of Olympic Games tends to be more positive.

7.2 Error Analysis

We can see form the figure that not every sum of S_{tk} is positive numbers, and some of the values is very big instead. This is because our data is not enough comprehensive. Because the cost of operating an Olympic spends not only the construction of the sports sites, but also the game operating fund, the facility upgrading fund and many other conjoint costs which contains a lot of items which is difficult to collect as different countries have different methods to impose these money. This part of money many be several times of the construction cost[3].

Another reason for the existence of the data outlier points is some emergency affairs makes the situation not in accordance with our assumptions. For example, the highest points in the figure is China. China, as a new host for the Olympic, spent large amount of money to newly build the sports sites.And there maybe another emergency things which disturb our model.

7.3 Innovative Suggestions

According to the conclusion that we get from the scatter chart, when *t* is near 0 point the total effect of Olympic Games tends to be more positive, we can predict that if the interval of Olympics can be shortened, the effect of Olympics on the hosting city will be more positive, indicating increasing willingness for cities to hold Olympic Games.

To achieve this, we suggest that the current Olympic mode can be divided into more periods and reducing the total cost of each event. One way is to divide the Summer and Winter Olympics into four seasonal parts and the other way is to separate the current Olympic venues into more cities so that each city can spend less on the construction fee.

In addition to this direct conclusion drawn from the scatter chart, we also find that there is some relationship between *S(t)* and the Greenhouse emissions of each country7. For instance, cities like Athens(in Greece) and Sochi(in Russia) have both low values of greenhouse emissions and *S(t)*, indicating more attention paid on environmental protection business can lead to a more positive comprehensive Olympic effect on the hosting city. As a result, we suggest that the more well-performed eco-friendly hosting forms will lead to more affection and support from common people and thus improving the city fame.

8 Strength and Weakness

8.1 Strength

- Our model has used the method of **Divide-and-conquer algorithm**, dividing the total question into similar smaller questions, and then solving the questions step by step. In this way, we can simplify the question effectively.
- Our data screening both focus on the quality of the data and the quantity of the data. For example, we abandoned the data after 2019, because of the impact of COVID-19 pandemic.
- We have used data from different sources and approaches. We used **crawler technique** to get the number of appearances of the news that was related to different cities and countries in the online news website. In this way, we effectively find the indicator that helps quantify the prestige of the city.

8.2 Weakness

- Our model lacks enough logical reasoning and theoretical deduction, so that we rely too much on data analysis. But the data we use is restricted to many realistic factors.This may make the calculation results reverse from the idealised condition we supposed before.
- Our model can adapt to the real situation when no dramatic events happen, but it can be adapted to the current case which the Covid-19 epidemic affects the world economic on a large scale.
- Our model analyze different indexes separately and regard them as a function of the only variable. However, in the reality, these indexes and the functions of them are all not truly independent. For example, we consider the urban development room as an important component of the city potential. However, as the land use is denoted by the GDP per unit of land area of the city, it is actually highly related to the economy of the city.This make our model failed to predict the impacts in the long run.

9 Conclusion and Further Discussion

9.1 Conclusion

In response to the fact that even fewer cities are willing to host the Olympic Games, we first use keyword search to identify that economy, city potential, city prestige are the three primary factors that we need to consider and draw a cloud figure to show their frequency of appearance together. We then collect data on secondary factors in question and normalized their characteristics by **feature scaling** in order to derive more accurately their quantitative relationships after being influenced by the Olympic Games with the strategy **polynomial curve fitting**. Based on these conclusions, we use the **Entropy Weight Method** to identify how much each of the primary factor accounts for the effect of Olympic Games on the specific city and construct our IC-Model matrix to predict what kind of modification of current Olympic policies can improve the Olympic effects on the city. With all these prepared, we finally plot a scatter chart to show the Olympic effect situations and make predictions of diverse changes of indicators based on the analysis of the scatter chart. With these predictions, we can then provide some proposals to IOC to improve the number of bids to hold Olympic Games.

9.2 Discussion

- There are some other precedent researches whose conclusions are in line with our conclusions. As Madden's article in 2006, he utilised the Input-Output Method to analyse the economic impacts of the Sydney 2000 Olympic on Australia's economy. He concluded that mega sporting events, even of the size of the Olympics, are unlikely to generate large economic benefits, but will increase the financial burden of the state government[8]. In Liu, D's 2017 research, all the indexes about social and environment was improved by the holding of the Olympic.[1] These researches support our own research and demonstrate that our results are reasonable.
- The holding of the Olympic is a cause that need long-term inputs that include lots of things like the sports site constructions, operating funds, the upgrade of public transportation,etc. But the host city will only be benefited form few aspects such as tourism, which is mainly concentrated

in the two-week period, and the following 3-4 years. Therefore, we can know that the impact on economy is negative. With the improvement of the appearance in the press, the city prestige will surely rise than before.[10]

- The model in this article is idealised because we consider each index we choose as a function of a very variable t. But in reality, these functions themselves are dependent to each other. Next we can utilise more complicated mathematical tool to make our research more precise.

10 Validation and Sensitivity Analysis

10.1 Verify the feasibility of our Model

To popularize and promote the IC-Model, it is important to test the validation and feasibility of our IC-Model. However, due to the small data set, we find it difficult to classify it proportionately into test set, training set and cross-validation set. So instead of categorizing, we choose one smaller group of real city inside the data set and calculated the economy, city potential and city prestige of it to figure out the weight matrix and compare it with the matrix we get in task 2.

The city we consider using to test our model include Vancouver, London, ... The rationale of our city selection is that the development of the economy of these cities is rather stable years before the Olympics. Meanwhile, unlike cities like Brazil and Korea, there were no major events during the years we evaluate.

By inputting the data of the feature the Vancouver during the Olympics, we were able to calculate the three indicators of different years of the cities we chose. Then we could use the Entropy Weight Method to get the weighted matrix according to our validation set.

$$w_t = \begin{bmatrix} 0.6107 & 0.2397 & 0.1496 \end{bmatrix}$$

To measure the similarity of the two matrix, we use **Enclidean Distance**, the real distance between the two dots that the one-dimensional matrix corresponds to in the three-dimensional space, to measure the similarity of the two matrix. we compute the Enclidean Distance between the weighted matrices we get from our test set and the whole data set respectively to verify the fitness of our model by computing the Enclidean Distance between the weighted matrices.

$$E(\omega_t - \omega_\omega) = \sqrt{\sum_{i=1}^{3} (\omega_{ti} - \omega_{\omega i})^2}$$

The weighted matrix we got from the whole data set before, w_w = [0.7758 0.0902 0.1340]. The weighted matrix we got from our test data set, w_t = [0.6107 0.2397 0.1496]. The Enclidean Distance between them is 0.2232, so we consider our model is valid within a range.

10.2 Analyze the Sensitivity

After building the model, it's important to test the sensitivity of the model to the different variables. By changing different variables while keeping others the same, we are enabled to investigate how

our IC-Models are affected by the characteristics of different models and also test the stability of our model.

In our IC-Model, the overall indicator, S_{tk} is a function related only to the year t, and the characteristic of city k. So to test the sensitivity of our model, we fixed the GDP of the ideal city and changed tourism revenues of the front year of holding the Olympics to investigate its influence. In the similar way, we kept the tourism revenues of the front year unchanged while testing the models with regard to different GDP of the city. To compare the two factors, we changed the tourism revenues and the GDP by 1 $\%$ each time. By plotting the graph of our model, the correlation between the feature of the city with the model can be clearly seen.

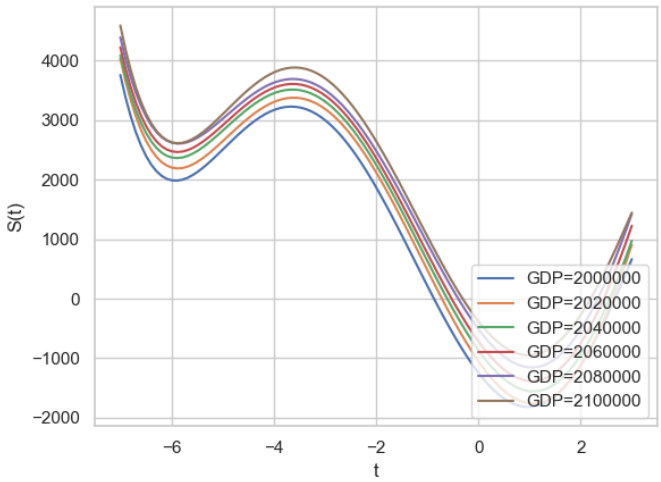

Figure 10: IC Models with changed GDP

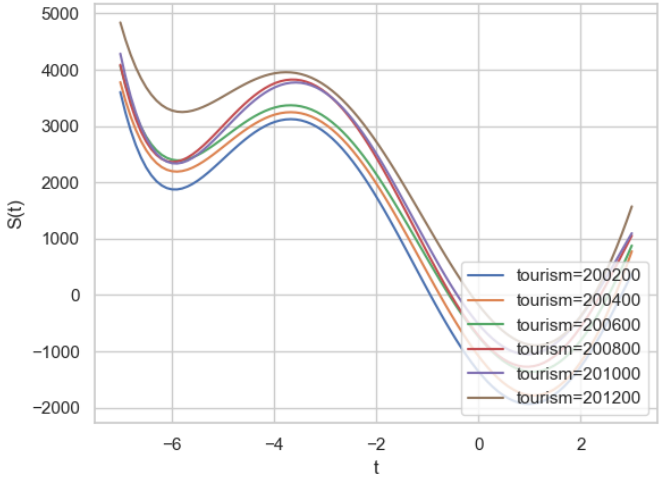

Figure 11: IC Model with changed tourism revenues

As was illustrated by the two graphs, the increase of both GDP and tourism revenues of the front

year of the city when the Olympics was held will cause the overall indicator become only slightly higher, which verified the stability of our model. What's more, in the comparison of the two graphs, we found that tourism revenues of the city before holding the Olympic was generally the factor that had more influence on the model than the GDP.

References

[1] Liu, D., Hautbois, C., Desbordes, M. (2017).The expected social impact of the winter olympic games and the attitudes of non-host residents toward bidding: The beijing 2022 bid case study. International Journal of Sports Marketing Sponsorship, 18(4), 330-346. doi:https://doi.org/10.1108/IJSMS-11-2017-099

[2] John R. Madden (2002) The Economic Consequences of the Sydney Olympics: The CREA/Arthur Andersen Study, Current Issues in Tourism, 5:1, 7-21, DOI: 10.1080/13683500208667904

[3] ShiNa Li, Adam Blake, Rhodri Thomas. Modelling the economic impact of sports events: The case of the Beijing Olympics

[4] Gulijiazi Yeerkenbieke, Chunci Chen and Guizhen He. Public Perceived Effects of 2022 Winter Olympics on Host City Sustainability.

[5] STEPHEN ESSEX * BRIAN CHALKLEY (2004) Mega-sporting events in urban and regional policy: a history of the Winter Olympics, Planning Perspectives, 19:2, 201-204, DOI: 10.1080/0266543042000192475

[6] Evangelia Kasimati*Department of Economics and International Development, University of Bath, Claverton Down, Bath BA2 7AY, UK. Economic Aspects and the Summer Olympics: a Review of Related Research.Int. J. Tourism Res. 5, 433–444 (2003) Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jtr.449

[7] Gordon Waitt University of Wollongong, Australia. SOCIAL IMPACTS OF THE SYDNEY OLYMPICS. Annals of Tourism Research, Vol. 30, No. 1, pp. 194–215, 2003.

[8] Madden, J. R. (2006). ECONOMIC AND FISCAL IMPACTS OF MEGA SPORTING EVENTS: A GENERAL EQUILIBRIUM ASSESSMENT. Public Finance and Management, 6(3), 346-394. Retrieved from https://www.proquest.com/scholarly-journals/economic-fiscal-impacts-mega-sporting-events/docview/237018349/se-2

[9] Naomi Kirkup Bridget Major (2006) Doctoral Foundation Paper: The Reliability of Economic Impact Studies of the Olympic Games: A Post-Games Study of Sydney 2000 and Considerations for London 2012, Journal of Sport Tourism, 11:3-4, 275-296, DOI: 10.1080/14775080701400943

[10] Baade, R.A. and Matheson, V.A. (2002), 'Bidding for the Olympics: Fool's Gold?', in Barros, C., Ibrahimo, M. and Szymanski, S. (eds), Transatlantic Sport: The Comparative Economics of North American and European Sports. Edward Elgar, Cheltenham; 127–51.

NEW WAYS FOR OLYMPICS

To: The International Olympic Committee (IOC)

From: Group#2333779

Date: April 3, 2023

Subject: Suggestions on the form of the Olympic Games

Nowadays fewer cities may be interested in hosting the Olympic Games. In order to change this situation and continue to spread the Olympic spirit, we have taken into account the time, the economy, the reputation of the city and the potential of the city (including environmental and land use factors) to come up with some suggestions that could help continue to host the Games

The Olympic Games are a significant event that brings together athletes, fans, and countries from around the world. However, hosting the Olympics can be a challenging and costly endeavor. To address these challenges, there have been several proposals for improving the hosting of the Olympic Games based on our mathematical IC-model. The main factors considered in our model can be classified as "hard power", like economy, and "soft power", like the city prestige to pluralize the evaluation level of our model on Olympics.

Our proposals are following:

-Adjusting the Interval between Olympic Games

The current interval of four years between Olympic Games has been in place since 1924. Shortening the interval could increase public interest and excitement in the Games and thus bringing a tourism boom. According to our analysis of how hosting Olympics will affect a city with IC-Model, we find that in the short period around Olympics the sum of its effects is usually positive and as time passes, these benefits decrease and the general effects go negative. Therefore, we suggest dividing current Summer and Winter Olympics into more periods and reducing

-Pay more attention to being eco-friendly

Based on our analysis on the data, when the greenhouse gas emissions of the hosting country during the Olympic time is low, it is more likely for the evaluation indicator of the city to be positive in a comparatively long period. We suppose this is mainly because more well-performed eco-friendly hosting forms will lead to more affection and support from common people and thus improving the city fame.