Problem Chosen	2024 HiMCM/MidMCM Summary Sheet	Team Control Number
A		14993

As the 2024 Paris Olympics have closed, the 2028 Los Angeles Olympics is on its way, and as the 2032 Brisbane Olympics are being meticulously planned, new sports, disciplines, and events (SDEs) have to be chosen to see what will be played at the largest sporting event in the world. These sports not only have to be popular and international, but also have to fit in with the International Olympic Committee's (IOC) various criteria, such as inclusivity and relevance. This means the sports chosen for the Olympics have to be carefully chosen.

Our model, which determines how well a SDE will perform, is a multivariable linear regression model with several interaction terms, the terms being the criteria required by the IOC quantified. These criteria are whether or not the sport has a governing body, which affects the sport's organization, its male to female player ratio, which affects its inclusivity, its revenue and web traffic, both of which indicate popularity, and year of inclusion as well as time in years.

Based on these variables, we have chosen three suitable SDEs to be introduced and/or reintroduced into the 2032 Brisbane Olympics. These three sports are cricket, the popular batting sport played by millions around the world, lacrosse, the American sport becoming increasingly popular, with Native American roots, and muay Thai, one of the most popular martial arts in the world.

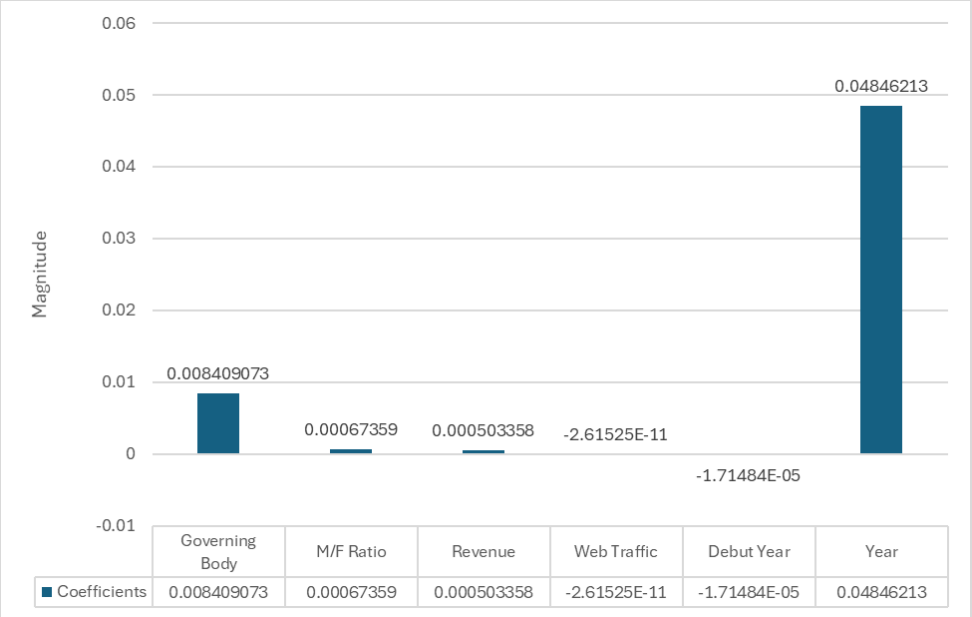

Figure 1: Coefficients of the Linear Regression

Contents

1 Introduction

1.1 Background Information

The Olympics is a momentous sport competition that is lauded by a global audience. It is characterized by its unique quadrennial element —meaning that it occurs once every 4 years—, its alternating locations, and its expansive range of sports that welcome representative athletes from diverse backgrounds.

Since its introduction in Greece in 1896, the Olympics has been held 30 times across a total of 21 host countries. Only 5 sports have been present in all iterations of the Olympics since the first one: Athletics, Cycling, Fencing, Gymnastics, and Swimming. Throughout the years, over 30 new events have been added. 14 sports were discontinued, 4 of which later revived through popularity vote while the remaining have been permanently discarded. The amount of events in each iteration of the Olympics have fluctuated, with 32 in the most recent 2024 Paris Olympics.

The decision of which sports are included is largely dependent on the International Olympics Committee (IOC), which comprises 111 members as of now. Internal voting is conducted on which events are added or discontinued. Apart from the IOC, the host countries have also played a role in this process since 2020.

Looking forward, individual sports, disciplines, and events (SDEs) are cemented for the upcoming 2028 Los Angeles Olympics. The IOC is currently redesigning SDEs that will reappear or disappear in the 2032 Olympics, which will be held in Brisbane.

As the Olympics is one of the most prestigious, largest grossing sport competition, the SDEs that are made available in each iteration holds paramount significance. Not only does selecting the SDEs impact us economically and culturally, it also affects athletes, coaches, and largely the future of certain events in the sport world.

In a case example, Jeu de Paume, one of the 10 discontinued sports, is now regarded as “jurassic tennis” and is replaced by more current sports (ie. tennis, squash). Although Jeu de Paume in its own right is still popular, its removal from the Olympics most definitely played a role in diminishing its once global renown.

Therefore, our team seeks to mathematically formulate an adaptive system to aid the IOC in its decision making process by providing a more holistic, analytical view on each potential SDE.

To aid the selection process of SDEs that will be added or removed in the upcoming 2032 Brisbane Olympics, our team must create a mathematical model that can quantify and weigh the advantages and disadvantages to each change made to the current SDE list.

1.2 Problem Restatement

To aid the selection process of SDEs that will be added or removed in the upcoming 2032 Brisbane Olympics, our team must create a mathematical model that can quantify and weigh the advantages and disadvantages to each change made to the current SDE list.

2 General Assumptions & Justifications

- **The IOC's commitment to values like sustainability and inclusivity will remain consistent through 2032.** Therefore, the mathematical model will assume that the evaluation criteria will not shift, allowing for a stable framework without needing to account for potential changes in these values. [21]
- **The infrastructural and technological advancements will continue to reduce cost by increasing accessibility and safety within sports.** We can expect reasonable costs and logistical requirements for new sports, assuming that the introduction of sports will have ongoing improvements. [17]
- **Data for more recent Olympic sports are available and will provide reliable predictions for future trends.** Historical data would be used to get better insight, as this method is practical and achievable to model. [3]
- **The demography across the viewers and participants, including interest in digital and new sports among the youth will continue.** Therefore, the model assumes the continued interest from the audience and uses this as a basis for providing a more accurate result in the relative relevance of various sports. [8]
- **The countries currently participating in the Olympics will continue to participate through 2032, with a minor impact on participation rates related to new sports.** Therefore the model assumes that the Games will have a stable reach and inclusivity, so it will not need to account for major geopolitical changes. [1]
- **The 2032 Olympics will maintain a similar scheduling and venue capacity compared to the recent Summer Olympics.** Therefore the model is going to assume the event schedule and venue capability are stable, enabling projections of logistical impacts realistically. [20]

3 Variables

Variable	Definition
a	Coefficient for Governing Body
b	Coefficient for Male-Female Ratio
c	Coefficient for Revenue
d	Coefficient for Year of Inclusion
e	Coefficient for Popularity
f	Coefficient for Time
g	Intercept
x_1	Governing Body (0,1 representing Doesn't exist/Exists)
x_2	Male to Female Ratio (Percentage)
x_3	Revenue (Category)
x_4	Year of Inclusion
x_5	Popularity
x_6	Time (Years)

4 Factors

We have defined variables x_1 to x_6 as factors that are to be considered in our model. From the given criteria provided by the IOC, we identified 6 elements that can be quantified or qualitatively analyzed: governing body as a boolean (Yes/No), male to female ratio as a percentage, revenue categorized by rank, year of inclusivity, popularity as a counted number, and time in years.

The majority of SDEs have an external organization (meaning it is not the IOC) that officiates rules and regulations within the sport that must be followed by its players. For instance, most aquatic sports like swimming fall under the institution "World Aquatics". The existence of an organization or governing body for this particular SDE is defined in our model as x_1. Another element emphasized across all sports and physical activities is safety. Behind fostering a healthy and secure environment for all athletes is the guarantee that the SDE itself is considered safe. This includes the ability to establish durable facilities and ensure no life threatening harm in more aggressive or combat SDEs in order to prevent athletes from sustaining permanent injuries or trauma. Often, the stricter and sport-specific regulations set by larger organizations prohibits athletes from performing riskier actions and enforces consistent audits on equipment and facility.

The number of male and female athletes was totaled for each individual year and computed as a percentage. A more equal male to female ratio, which is defined as x_2 in our model, is prioritized because equality across SDEs is a key objective of the IOC among many others.

Further research into the revenue provided us with an objective ranking of the profits garnered in each of the 26 SDEs in the 2016 Olympics in Rio de Janeiro. The original reports categorize the revenue from Groups A to E, thus we established 5 tiers in our own model, with which we will rank the SDEs. As organizing and coordinating the Olympics is not only strenuous, but also requires extreme resources (given that the location changes every iteration and each SDE necessitates professional, large scaled equipment), it is crucial for the IOC and the host country to be able to generate a profit from holding the competition. In our scale, we categorize SDEs from 1 to 5 using variable x_3, with a ranking of 1 allowing for the least amount of profit and a ranking of 5 for the most.

Our variable x_4 represents the year at which the SDE was included in the model. In other words, this variable determines the duration for which the sport has been available as an event that athletes can compete in. This variable allows the model to determine whether the SDE has been consistently included in each iteration of the Olympics or if it has been discontinued before.

In order to maximize the revenue that is described above, the popularity of the SDE needs to be considered. The number of athletes in each SDE, which we have scaled down in the variable x_5, suggests both the amount of recognition of each sport in the athletic world and the competitiveness of each sport too, with a higher number of participants decreasing the chances of performing well in the Olympics.

Finally, time is the final variable that we integrated into our model is time in years, which is represented by x_6. This variable allows us to specify the iteration of the Olympics that is the model is computing for and it enables the system to use predetermined trends of the other factors to extrapolate data that can be used in the selection process of the SDEs.

The variables a to f are all respective coefficients of the variables x_1 to x_6. They enable the model to weigh each variable's effect on the final score for the SDE differently. Through linear regression, we are able to train these coefficients to maximize the accuracy of our model and improve the reliability of our extrapolated trends in relation to actual trends.

5 Model

5.1 Model Inputs & Outputs

The factors that determine an Olympic sport being added or removed range from easily measurable factors, such as the popularity, to less measurable factors such as the gender equality. As such, we need to be able to quantify different the different inputs and outputs. Additionally, the model needs to be able to evaluate different trends into the future, so the time needs to be considered as well. Each input will be individually passed into the model and the final output will be a scalar weight that is proportional to the number of events the sport is expected to have. The weight represents how likely an Olympic sport is to succeed. The purpose of this is so that the model can utilize the past Olympics to train itself. In summary, the model will input in the independent variables, then output a final weight which is the model's decision on how likely the sport is to succeed in the Olympics that year.

5.2 Model Design

As the Olympics is a varied competition and has a wide range of sports, our model needs to have low **variance** (not over fit the training data, causing it to model the noise) yet have as low of a **bias**, or error, as possible. Since the model needs to be able to select a variety of SDEs, variance is preferred. For this reason, the model selected is a multivariate linear regression model with interaction terms. The general form of a multivariate linear regression is $y = ax_1 + bx_2 + cx_3 ... nx_n + intercept$ where a, b, c, etc. are coefficients, and $x_1 ... x_n$ are independent variables. An intercept is also generated, representing the control environment. If no independent variables are used, the standard weight of a SDE would be the intercept. A multivariate linear regression identifies the best coefficients, which can be determined through a multitude of methods, mathematical and heuristic, that give the best predictions of y - the dependent variable. In this case, the dependent variable chosen during training was the number of events the sport had in any given year. In order to evaluate the model, the predicted values are compared to the actual values, and the squared difference is summed up to create a sum of squares residual.

However, as time plays an important role in the selection of Olympic sports, through trends and number of sports, this multivariate linear regression is not able to completely model time's effects on the Olympics. As such, interaction terms were added to allow the model to change it's coefficients as time progresses. The equation for a multivariate linear regression with interaction terms is $y = ax_1 bx_2 + cx_3 bx_4 ... nx_n + ox_(n + 1)$, and any variables can be multiplied with each other. In this case, since time is the interacting term, we multiply all of the terms by time except for the intercept and time. The equation we use is $y = ax_1x_6 + bx_2x_6 + cx_3x_6 + dx_4x_6 + ex_5x_6 + fx_6 + g$ Where y is the weight, x_n is an independent variable, and a to g are coefficients. You can see that for every term, as x_6 - the time - changes, the resulting term also changes depending on the sign of the term. If the coefficient is large, for every change in x_6 there will be a corresponding large change in y. The sign and magnitude of the coefficient determines the trend of that specific independent variable: a negative sign indicates that the independent variable contributes negatively to the events, so a lower value of that independent variable is preferred. The magnitude indicates the importance of the variable, a higher magnitude means that for every increase in x_n there is a corresponding greater increase. Factoring out x_6, the final equation is $y = x_6(ax_1 + bx_2 + cx_3 + dx_4 + ex_5 + f) + g$.

5.3 Model Training

The coefficients determine the relative importance of each independent variable and the final equation. After we determine the final coefficients, we can determine the final equation, which we can then use to evaluate new SDEs. As such, determining the coefficients is the basis of the model.

The data we use to train the model is the past data of events from past Olympics. For every year and discipline, we create a prediction of how many events there were through the regression. However, there will be some error from the actual number of events. Since this error could be negative or positive, the error is squared in order to ensure that the error is positive. The squared error is the metric that is used to evaluate the model.

In order to train the model, we identify the coefficients that minimize the squared error. An efficient yet simple way to minimize the squared error is coordinate descent [4]. Through this method, the model optimizes one coefficient at a time then continuously cycles through until the improvement is minimal, to reduce overfitting. Coordinate descent has the benefit of efficiently optimizing each individual coefficient. Additionally, it also works with interaction terms, as there is no cost function needed - just the sum of squared error. Since the coefficients are extremely important to the model, the coefficient-centric coordinate descent works very well with our model.

To evaluate the model objectively, R-squared is used as a metric as well, comparing the sum of squared error with the error one would get by using the average only. The r-squared shows how much of the error is explained - a r-squared of 0 indicates that the model is no better than guessing the average, and a r-squared of 0 indicates a perfect model.

The final coefficients are: 0.008409073, 0.00067359, 0.000503358, -2.61525E-11 -1.71484E-05,0.04846213, as seen in Figure 2 with a R-squared of 0.288. While is the r-squared is quite low, it is able to explain much of the error and is significantly better than using the average, suggesting that it is able to model some trends. The intercept generated was -46.2411, which was not shown as it does not help predict the event's weight.

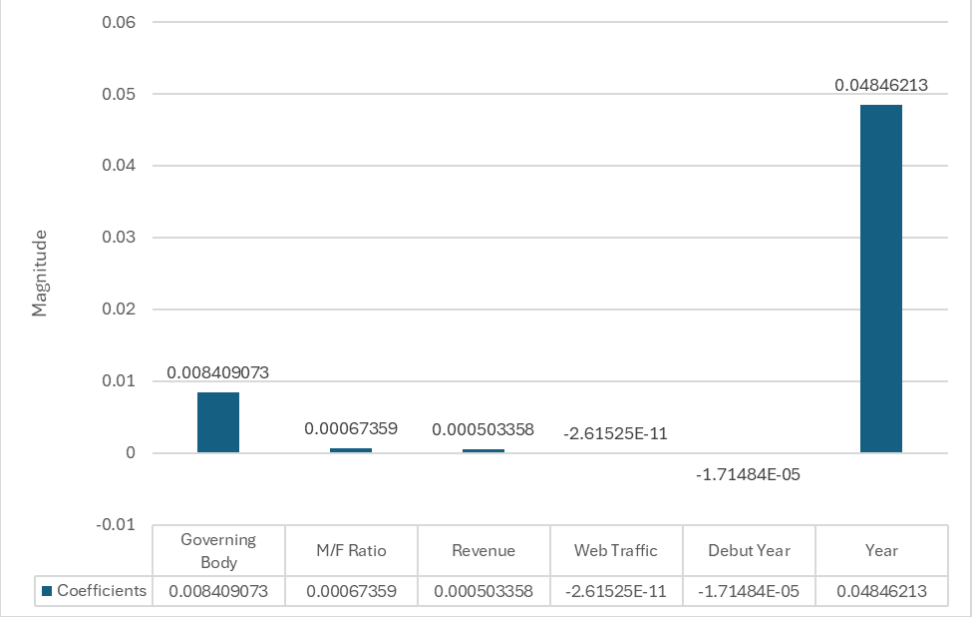

Figure 2: Coefficients of the Linear Regression

The magnitude of each coefficient corresponds to the importance of each variable. Since the intercept is -46.2411, the initial prediction for any sports is -46.2411, which is good for predicting new SDEs, as we are more concerned with having an extremely strong weight, and having a very relevant Olympic sport.

5.4 Model Usage

The entire linear regression model's main job is to predict the future events of a specific SDE. The coefficients that are trained merely create an equation that we can use to predict the success - measured by how many events a sport is expected to generate - of a SDE. These coefficients are also analyzed further under 7.

To use the model, we first identify the independent variables of a new specific SDE. Then, after identifying the independent variables, we then substitute the independent variables into the general equation. The resulting value will be a weight that represents the expected number of events, with a greater weight signifying a higher likelihood that the sport will be continued into further Olympics.

6 Potential SDEs

6.1 Cricket

Cricket should be introduced into the Olympics as an SDE for the 2032 Olympic Games

A sport that should be reintroduced at the 2032 Olympic Games at Brisbane should be Cricket. Utilizing a shortened style of cricket in the cricketing world, T20, [22] Cricket is a perfect SDE to add to the Olympic Games, fitting all 6 Olympic criteria too.

Cricket undoubtedly fits the first criteria for the Olympic committee, popularity and accessibility. Being extremely popular in India and Pakistan, both of which are developing nations with an increasingly substantial population with access to the internet and therefore Olympic coverage, Cricket has the backing of a large population that would propel cricket into a major Olympic sport. This rising interest can be shown through an analysis of web traffic, and cricket, with 64.4 million average searches monthly from November 2023 to November 2024, and an average of 55.6 million of those searches being from India, showing the potential of cricket as a prevalent sport. [11] The overall predicted revenue generated within the cricket market is around 3.71 billion USD in 2024 [5], similar to the 4.01 billion USD generated by rugby in 2023. [16] As rugby was given a 1.0 on our revenue scale based on the IOC's allocation of funding, we can assume that cricket will generally have a similar allocation of funding and generated amount of revenue within the Olympics, and has a 1.0 on the revenue variable, x_3, making it relatively profitable and popular.

Similarly, it is very accessible, with few necessary resources needed to play, as bats, protective gear, wickets, and a field are the only necessary requirements to start a game of cricket. [2]

As for gender equity, as of 2017, only around 10% of cricket players are female. [15] This gives cricket a 0.9 on the gender equity variable, x_2, which means most players are male. But with enough support that proportion could be 50-50; there is also no significant obstacle for female players to play cricket, ensuring that both genders have equal opportunities to participate in cricket.

Cricket follows the basic guidelines of environmental and social responsibility as much as other sports. The sport in itself does not produce any carbon emissions, unlike sports such as auto racing, and

the gameplay of the sport does not include any sort of action that might incite social irresponsibility.

Cricket also fits the requirements for inclusivity. There are 108 member nations within the International Council of Cricket, 12 full members and 96 associate members, hailing from all continents except Antarctica. This fits the 75 countries from 4 different continents requirement set by the International Olympic Committee and shows the global spread of cricket. [18]

Cricket is also a sport of innovation and relevance, not just because of its popularity, but also because of recent innovations to the variations of cricket, most notably the T20 variation, which gives cricket games a shorter game time, allowing for the game to be more accessible, easy to watch, and fitting with Olympic schedules. Nevertheless, the traditional rules of cricket still remain, allowing cricket to be a sport that blends innovation with tradition.

Like many other sports, cricket is committed to safety and fair play; doping is not excused in this sport more than any other sport, and cricket players wear protective gear to ensure the safety of the players. The existence of an international committee, indicated by the variable x_1, shows that there is a governing body regulating the sports as a whole.

6.2 Lacrosse

Another sport that should be reintroduced at the 2032 Olympic Games in Brisbane should be lacrosse. From our model, we predict lacrosse to be the best out of all 3 potential SDEs we evaluated and we encourgae you to supplement it to the 2032 Olympics. Similar to cricket, it fits all 6 Olympic criteria.

Lacrosse is an increasingly popular sport, fitting the Olympic popularity and accessibility. Being increasingly popular in the United States, analysis of web traffic shows 389,000 average searches monthly from November 2023 to November 2024, comparable to archery with 301,900 average searches. It is also very accessible, with only a field, stick, and protective gear being needed for most players. [6] On the revenue side, lacrosse had a market size of $312.5 million USD in 2023 [12], sizable to table tennis's $485.3 million USD [19], landing both sports with a 3.0 on the revenue variable, x_3, using the criteria described above.

As for gender equity, as of 2021, only around 18% of professional lacrosse players are female. [7] This makes the gender equity variable, x_2, 0.82. However, similar to cricket, there is no significant hindrance stopping the number of female lacrosse players from increasing, and lacrosse can be made accessible to all genders.

Lacrosse follows environmental and social responsibility as much as other sports. The sport in itself does not produce any carbon emissions, and the gameplay of the sport is not violent and does not include any sort of action that might incite social irresponsibility.

Lacrosse also fits the requirements for inclusivity. World Lacrosse has 92 member nations over 4 continents playing lacrosse, [23] which fits the 75 countries from 4 different continents requirement set by the International Olympic Committee and shows the global spread of cricket. Lacrosse's roots in Native American culture also show that its roots are inclusive, as it has origins from many different cultures.

Lacrosse is also a hallmark of innovation and relevance. Unlike many other sports where the technology used in the sport has staggered, lacrosse sticks continue to be technologically innovative and progressive. Culturally, its Native American roots bring it cultural relevance as modern society tries to remember and celebrate the impact of Native cultures around the world.

Like many other sports, lacrosse is also committed to safety and fair play. Protective gear is used by all players, and any sort of drug use or doping is not at all allowed within lacrosse. Since it also has

a governing body regulating rules within the sport, it gets a 1 on for the variable x_1.

6.3 Muay Thai

The last sport that should be reintroduced at the 2032 Olympic Games at Brisbane should be Muay Thai.

Muay Thai is an extremely popular sport; analysis of web traffic shows that there are an average of 1,000,000 searches monthly from November 2023 to November 2024. This is comparable to many other big sports within the Olympics, such as Baseball and Taekwondo. It is also very international, as only around 165,000, or 16.5% of those searches, are American. Muay Thai gyms can be seen everywhere around the world, and cost-wise, not much funding is needed, other than a ring and gear. Muay Thai also has a $1.25 billion USD market, an indication of its popularity. [13] This is similar to ice hockey's $1.9 billion [9], putting it at a 2.0 on our revenue variable, x_3, showing that it is quite profitable.

The IFMA, or the International Federation of Muaythai Associations, claims that many Muay Thai sporting events have an equal amount of male and female competitors, such as the ones at the European Games. This makes Muay Thai truly gender equitable,[14], and it also gives Muay Thai a 0.5, a perfect score, on the gender equity variable, x_2. This suits the Olympic criteria perfectly.

Muay Thai, while violent, is not unorganized, and has set, civilized rules, meaning it does not offset social responsibility in any way. This is similar to many other combat sports within the Olympics, such as boxing and taekwondo, and Muay Thai would perfectly fit into the growing list of martial arts within the Olympics. Like many other sports, it also respects the guidelines of environmental responsibility.

IFMA has 146 member nations spanning all continents but Antarctica, which makes Muay Thai an international sport and fitting of the criteria set by the IOC.[10]

Muay Thai's increasing popularity around the world shows its relevance in modern sports today. How the fighting rules and the scoring system have changed shows Muay Thai's commitment to innovation. These rules are fluid, meaning they can be adapted to appeal to a larger audience, such as those of the Olympics.

While safety is a major concern within Muay Thai games, like all other martial arts within the Olympics, regulative and restrictive measures should be placed, and protective gear should be properly given, as well as proper refereeing. Doping is also not allowed within Muay Thai whatsoever. The IFMA is also Muay Thai's governing body, giving it a 1 on the governing body variable x_1. This shows that Muay Thai's regulations are enforced.

6.4 Model Prediction

After collecting data on the SDEs, we can use our model to predict the expected number of events. The independent variables of our proposed SDEs have been described above. Substituting these variables into the model returns the predicted number of events in 2032.

As you can see in Figure 3, all of the proposed SDEs are positive - there is no event that is predicted to have a negative number of events. This suggests that all of the proposed SDEs are fine, yet lacrosse is in the clear lead with a weight of 6.2739. This means that lacrosse is estimated to be the most valuable sport for the 2032 Olympics, and the weight suggests that lacrosse is an excellent sport

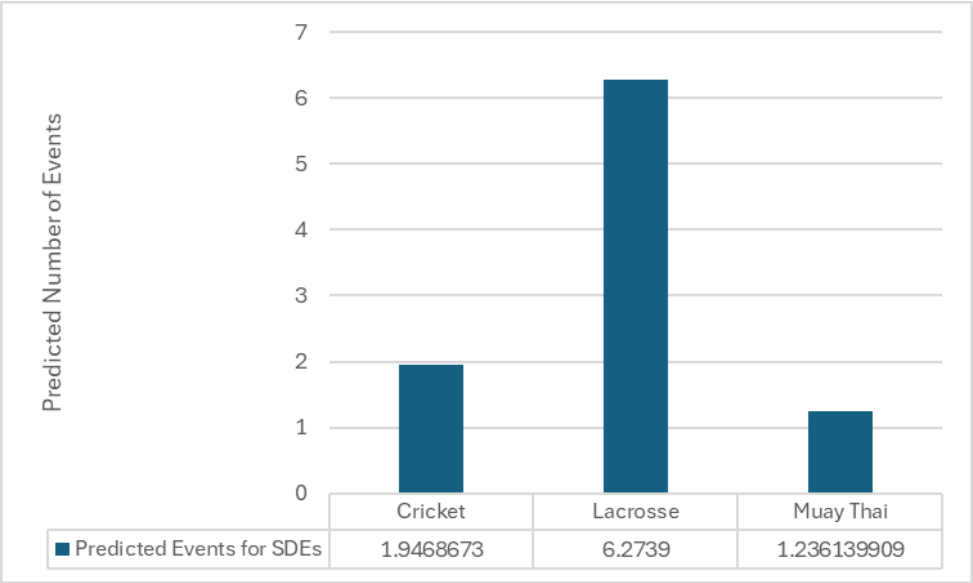

Figure 3: Predicted Events for Proposed SDEs

to add to the 2032 Olympics. However, cricket and Muay Thai are still fine additions, as both of them are expected to succeed.

7 Sensitivity Analysis

While we do have a metric for measuring how well the model predicts the data set - r-squared -it is unclear if the model is **statistically significant**. The relation may just be by chance and not any underlying trends or correlations. The claim that the fit is purely by chance is the null hypothesis. To validate this, we use a f-test to validate the statistical significance. Setting a significance value of p=0.05, our model clears the test: the F-distribution returned 2.3518e-162, an impressively low number. This means that our model is statistically significant as the chance that this fit occurred by chance is extremely low.

However, some of the coefficients in the model may be misleading: it seems that debut year and web traffic are extremely insignificant, but they just have much larger data - web traffic is in the 5,000,000s, and years range from 1984-2028. As such, it is much easier to see their relative effects after normalization. By multiplying each coefficient by the maximum output, the maximum output of each term can be identified. After regularization, the coefficients and their maximum effects can be identified in Figure 4.

It is important to analyze the coefficients and determine their significance. The current year is extremely important - there is an upwards trend every year, as the Olympics gets larger. As such, there are more events. However, not all events increase over time, as evidenced by the balancing factor of the debut year. Since the debut year is negative, we conclude that the lower the debut year, the greater the number of events. The older the sport has existed, it is more likely for it to have more events.

However, for new SDEs, their debut year and the current year will be the same across all new SDEs in a certain year. In the 2032 Brisbane Olympics, debuting SDEs will have the x_5 and x_6 be 2032. As such, what determines a new SDE's relative weighting in Brisbane 2032 are the other four

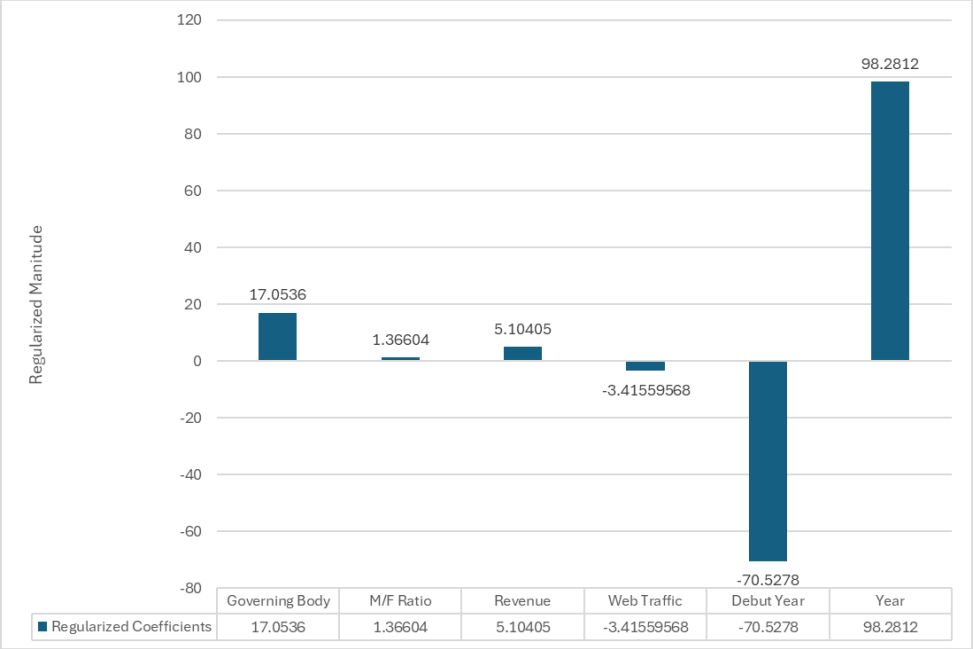

Figure 4: Regularized Coefficients

factors: If there is a governing body, male to female ratio, revenue, and web traffic.

Of these, the existence of a governing body is the most important - with a coefficient of 17.0536. Additionally, revenue is also quite important, with a coefficient of 5.10405. From these, we can deduce that a governing body is extremely important for a sport to succeed in future Olympics, and the revenue that it can create as an Olympic sport - it's competitiveness against other sports such as Athletics or Aquatics also plays a significant role.

Two outliers are the coefficient for web traffic, which is a negative. This suggests that a high web traffic is detrimental, however, it may also be overfitting to noise. Additionally, the coefficient for the male to female ratio is very low, which may suggest that it is unimportant.

However, it is important to note that while the coefficients do give valuable information, they only provide values which is used to **predict the success of Olympic SDEs**, which is the main goal of the model.

8 Letter

Letter to the International Olympic Committee.
To: Members of the International Olympic Committee
Subject: Formulating the Future of the 2032 Brisbane Olympic Games

The Members of the International Olympic Committee,

The Olympics are a test to the unlimited potential of human connection, resilience, and ambition. Every four years, the world joins together—not just to witness athletic greatness but to celebrate the spirit of humans that knows no borders, languages, or cultures.

While considering any selection of sports, disciplines, and events, one has to bear in mind that the Olympics are much more than competition. They are a field of dreams, a place where traditions

are followed, and new visions for the future are created. Each SDE chosen tells a story—one of hope, persistence, and dynamic evolution of world culture.

We strongly encourage you to use this opportunity to reflect not just the values of the present but also the hopes of the future. The right SDEs are those that will be presented—those that will amplify voices long silenced, give recognition to traditions so deserving, and build up passions for the next generation; those that consider sports that uplift, bring people together, and push the boundaries of what we think is possible.

With our model, the once subjective selective process is now marginally more holistic. The multivariate linear regression allows you to extrapolate key data from the past 30 iterations of the Olympics while the element of interaction terms allow the model to precisely visualize the effect of one variable on another. Similarly, our model is capable of adaptation when it receives new information, hence improving its accuracy.

Our bid is based on this vision. Cricket, lacrosse, and Muay Thai are not sports; they are passports to new communities, celebrations of inclusion, and reverences to creativity. In other words, adding these SDEs to the Olympic family sends a signal to the world that the Games are not limited by what has been but are truly dynamic, forward-looking, and universal.

Imagine the fields of Brisbane reverberating with the cadence of cricket, echoing the cheers of billions from Asia, Oceania, and beyond. Envision lacrosse, whose origins are inextricably woven into indigenous traditions, bringing respect and opening dialogue for all sorts of cultures. Feel the strong and graceful flow of Muay Thai, embodying strength, respect, and the uncompromising attitude to achieve mastery.

Every choice, every decision counts. It forms the fabric of who we are as an international community and what we hope for in the years to come. Olympics are a force of change, reminding us that there are many, many more things that unite us than things that divide us. The Games are not about the athletes but the millions upon millions of hearts they touch.

And as you make all these critical decisions, we ask you to look beyond the statistics and logistics. Think about the stories to be told, the connections to be forged, and the dreams to be given flight.

Together, we can craft an Olympic program that speaks to the world, not just in 2032 but for generations to come.

I admire your dedication to this cause. We are prepared to offer our support in bringing this vision to fruition.

Sincerely,
Team 14993

9 Strengths and Limitations

The main strength of the model is the broadness and wide variety of situations it can adapt to and model. It is extremely flexible and can scale very easily. With a statistically significant fit, it is accurate and is able to use the independent variables to produce predictions with purpose.

Since we use interaction terms and use multivariate linear regression, a wide variety of interactions and variables can be modeled. Trends and weights of different independent variables can be modeled, which allows for the model to be extremely broad. Different trends as time increases are accounted for.

The model utilizes coordinate descent to learn, allowing it to optimize all of the coefficients, yet also allowing the model to be trained on a broad set of data. This also allows the model to utilize a wide variety of training data - this model can be trained on new data, after the future Olympics have concluded.

As previously said, the model is extremely scalable and flexible. This multivariate linear regression could have more terms added to it, allowing it to scale as the Olympics gets more complex. Additionally, different independent variables can be added or removed from the model, giving it additional flexibility.

The complete solution accounts for all of the IOC criteria [7] :

- Popularity and Accessibility: Web traffic, revenue generated are both factors used in the model itself.
- Gender Equity: quantified through the male-female ratio of athletes in every sport, also used in the model
- Sustainability & Inclusivity: proposed SDEs have been checked to ensure the meet the criteria
- Relevance and Innovation: Recent trends are modeled through interaction terms
- Safety and Fair Play: research has been conducted to ensure proposed SDEs have athlete protection and anti-doping regulations, and the governing body is weighted heavily by the model.

The main limitation of our model is the large error, as the Olympics is difficult to fully model as there are many factors that were not considered.

The model is limited in it's independent variables. While the basic needs are met, there could be highly significant independent variables left unaccounted for in this model. Additionally, only time is modeled as an interaction term. There could be more unmodeled interactions between variables.

The model has a high r-squared, which means that it has a high error. With a r-squared of 0.288, the model accounts for nearly 30% of the error, but there is still a large fraction left. The model is able to create semi-accurate predictions, but these predictions vary widely. It's only able to give "ballpark" range predictions, but it is not precise enough to give closer predictions.

While we can model a wide range of interactions with different variables and interaction terms, it is limited to linear relationships. This means that trends like quadratic or other polynomial trends can't be accounted for: an event might peak and then fall back down, but our model cannot model that.

10 Conclusion

Our model is constructed on the basis of quantitatively transforming and easing the selection process. This model is able to holistically consider a multitude of variables that are recognized by the IOC, which are listed on the criteria. It adapts to new data in order to extrapolate accurate results.

With the introduction of our model, the IOC will hopefully entertain additions of new or previously discontinued SDEs. Identifying a sport that can both garner high revenue while remaining safe and inclusive is difficult; but our model may be able to change the game for future iterations of the Olympics.

11 References

References

[1] Mike Weed et al. *The Olympic Games and raising sport participation: a systematic review of evidence and an interrogation of policy for a demonstration effect*. 2015. URL: https://www.easm.net/download/easm_essential_sport_management_collection/sport_governance_and_policy/The-Olympic-Games-and-raising-sport-participation-a-systematic-review-of-evidence-and-an-interrogation-of-policy-for-a-demonstration-effect.pdf.

[2] *All About Cricket Equipment*. URL: https://cricketdirect.co.uk/pages/all-about-cricket-equipment.

[3] V Asha. *Performance Analysis of Olympic Games using Data Analytic*. 2023. URL: https://ieeexplore.ieee.org/document/10084943.

[4] *Coordinate descent*. 2016. URL: https://www.stat.cmu.edu/~ryantibs/convexopt-S15/lectures/22-coord-desc.pdf.

[5] *Cricket - Worldwide*. 2024. URL: https://www.statista.com/outlook/amo/sports/cricket/worldwide#:~:text=Revenue%20in%20the%20Cricket%20market,US%242%2C275.00m%20in%202024)..

[6] *Equipment*. 2024. URL: https://www.usalacrosse.com/equipment.

[7] *Factsheet: The programme of the Games of the Olympiad*. 2024. URL: https://stillmed.olympics.com/media/Documents/Olympic-Games/Factsheets/The-programme-of-the-Games-of-the-Olympiad.pdf.

[8] *Future Trends in Youth Sports Participation*. 2024. URL: https://sportsvisit.org/2024/08/13/future-trends-in-youth-sports-participation/.

[9] *Ice Hockey Equipment Market Research Report By Product Type*. 2024. URL: https://www.marketresearchfuture.com/reports/ice-hockey-equipment-market-23509.

[10] *IFMA Family*. 2019. URL: https://muaythai.sport/ifma-family/.

[11] *Keyword Overview: Cricket*. URL: https://www.semrush.com/analytics/keywordoverview/?q=Cricket&db=us.

[12] *Lacrosse Equipment Market Report by Type*. URL: https://www.imarcgroup.com/lacrosse-equipment-market#:~:text=Market%20Overview%3A,5.67%25%20during%202024%2D2032..

[13] *Muay Thai takes the US by storm, with 1,724 centers,* $1.25 billion industry$. 2024. URL: https://www.nationthailand.com/thailand/general/40034892.

[14] *Muaythai Praised for Full Gender Equality at the European Games*. 2023. URL: https://muaythai.sport/muaythai-praised-for-full-gender-equality-at-the-european-games/.

[15] Paul Radley. *ICC Want to Achieve a 50-50 Split of New Male and Female Participants in Cricket by 2023*. 2017. URL: https://www.thenationalnews.com/sport/cricket/icc-want-to-achieve-a-50-50-split-of-new-male-and-female-participants-in-cricket-by-2023-1.625713.

[16] *Rugby Market Size, Share, and Trends Competitive Analysis by Type*. URL: https://www.futuredatastats.com/rugby-market#:~:text=What%20is%20the%20current%20size,USD%206.3%20billion%20by%202030..

[17] Erin Scheg. *How is Advanced Technology Improving the Olympics?* 2024. URL: https://www.c2st.org/how-is-advanced-technology-improving-the-olympics/.

[18] Mohak Singh. *How many countries in the world play cricket*. 2023. URL: https://www.zapcricket.com/blogs/newsroom/how-many-countries-in-the-world-play-cricket?srsltid=AfmBOoq_YjrfTC12Icru_Z_dU3YZjd_LYvxI3Kwnnye30Xi0aVyP6v3u.

[19] *Table Tennis Table Market Size, Share, Growth, and Industry Analysis, By Type*. 2023. URL: https://www.businessresearchinsights.com/market-reports/table-tennis-table-market-110915.

[20] Stephen Wade. *Tokyo Olympic venues lined up, schedules remain the same*. 2020. URL: https://www.sandiegouniontribune.com/2020/07/17/tokyo-olympic-venues-lined-up-schedule-remains-the-same/.

[21] Garrett Wong. *IOC Sustainability Report Highlights 2017-2020 Achievements Across The Olympic Movement*. 2022. URL: https://www.greensportsalliance.org/media/ioc-sustainability-report-highlights-2017-2020-achievements-across-the-olympic-movement.

[22] Robert Wood. *What is Twenty20 (T20) Cricket?*. 2009. URL: https://www.topendsports.com/sport/cricket/twenty20.htm.

[23] *World Lacrosse Members*. 2023. URL: https://worldlacrosse.sport/world-lacrosse-members/.