2025 MCM
Problem C: Models for Olympic Medal Tables

In addition to watching individual **events** during the latest summer Olympic Games in Paris, 2024, fans followed the overall "medal table" for each country. The final results (**Table 1**) showed the United States with the most total medals (126), and China and the United States tied at the top of the standings for the first place Gold medals (40). The host country, France, was 5th in the Gold medal count (16) standings but 4th in terms of total medal count while Great Britain, 7th with 14 Gold medals finished 3rd in total medals.

	Gold	Silver	Bronze	Total
United States	40	44	42	126
China	40	27	24	91
Japan	20	12	13	45
Australia	18	19	16	53
France	16	26	22	64
Netherlands	15	7	12	34
Great Britain	14	22	29	65

Table 1: Paris Olympics (2024) Final Medal Table – Gold Medal Top 7 Countries[1]

The standings at the top of the table are always watched closely, but the medal counts for other countries are often just as valued. For example, Albania (2 medals), Cabo Verde, Dominica, and Saint Lucia (2 medals) won their nations' first Olympic medals at the Paris games. Dominica and Saint Lucia also each earned a Gold medal at these games. More than 60 countries have still yet to win an Olympic medal.

Predictions of the final medal counts are commonly made, but typically not based on historical medal counts but closer to the start of an upcoming Olympic games when current athletes scheduled to compete are known (for example: https://www.nielsen.com/news-center/2024/virtual-medal-table-forecast/).

Data is provided of medal tables for all summer Olympic games, host countries, as well as the number of Olympic events at each games broken down by **sport** for all summer Olympic games played. Additionally, data for all individual Olympic competitors with their sport and result (medal type, or no medal) is provided. Your models and data analysis must ONLY use the provided data sets. You may use additional resources to provide background and context or help with interpreting results (be sure to document the sources). Specifically, use the provided data to:

- Develop a model for medal counts for each country (for Gold and total medals at a minimum). Include estimates of the uncertainty/precision of your model predictions and measures of how well model performs.
    - Based on your model, what are your projections for the medal table in the Los Angeles, USA summer Olympics in 2028? Include prediction intervals for all results. Which countries do you believe are most likely to improve? Which will do worse than in 2024?
    - Your model should include countries that have yet to earn medals; what is your projection for how many will earn their first medal in the next Olympics? What sort of odds do you give to this estimate?
    - Your model should also consider the events (number and types) at a given Olympics. Explore the relationship between the events and how many medals countries earn. What sports are most important for various countries? Why? How do the events chosen by the home country impact results?
- Athletes may compete for different countries, but it is not a simple matter for them to change due to citizenship requirements. Coaches, however, can easily move from one country to another as they do not need to be citizens to coach. There is, therefore, the possibility of a "great coach" effect. Two possible examples of this include Lang Ping[2], who coached volleyball teams from both the U.S. and China to championships, and the sometimes-controversial gymnastics coach, Béla Károlyi[3], who coached Romania and then the U.S. women's teams with great success. Examine the data for evidence of changes that might be due to a "great coach" effect. How much do you estimate such an effect contributes to medal counts? Choose three countries and identify sports where they should consider investing in a "great" coach and estimate that impact.
- What other original insight(s) about Olympic medal counts does your model reveal? Explain how these insight(s) can inform country Olympic committees.

Your PDF solution of no more than 25 total pages should include:
- One-page Summary Sheet.
- Table of Contents.
- Your complete solution.
- References list.
- AI Use Report (If used does not count toward the 25-page limit.)

Note: There is no specific required minimum page length for a complete MCM submission. You may use up to 25 total pages for all your solution work and any additional information you want to include (for example: drawings, diagrams, calculations, tables). Partial solutions are accepted. We permit the careful use of AI such as ChatGPT, although it is not necessary to create a solution to this problem. If you choose to utilize a generative AI, you must follow the COMAP AI use

[policy](). This will result in an additional AI use report that you must add to the end of your PDF solution file and does not count toward the 25 total page limit for your solution.

Data Files

2025_Problem_C_Data.zip: This zip file contains all 5 of the data files listed below.

- *data_dictionary.csv* – database descriptions with examples
- *summerOly_athletes.csv* – all competitors with their sport, year, and result (medal type or none)
- *summerOly_medal_counts.csv* – complete country medal count tables for all summer Olympics from 1896 to 2024
- *summerOly_hosts.csv* – list of host country for all summer Olympics from 1896 to 2032
- *summerOly_programs.csv* – counts of number of events by sport/**discipline** and total for all summer Olympics from 1896 to 2032

Data, such as country designations, are recorded by the **International Olympic Committee (IOC)** (on their Olympics.com website) at the time of a given Olympics. Thus, designations may change in the data set. As with all data, there may be recording anomalies. Note, for example, in the athlete's data set in some cases for sports like tennis, table tennis, beach volleyball, the "Team" includes more detail than just the country. For example, Germany-1 would be the first of two beach volleyball teams from Germany in the 2000 Olympics. Decisions and assumptions about how to handle the data are an important part of the modeling process.

Glossary

International Olympic Committee (IOC): is the international, non-governmental, sports governing body of the Olympic Games and the Olympic Movement. The IOC is best known as the organization responsible for organizing the Summer and Winter Olympics.

Programme: of the Olympic Games is the programme of all sports competitions established by the IOC for each edition of the Olympic Games.

SDE: **Sport, Discipline,** or **Event**

> **Sport**: The IOC defines an Olympic sport as a discipline that is governed by a single international sports federation (IF). A single sport may contain one or more disciplines, each of which is the focus of one or more events.

Discipline: A branch of a sport that includes one or more events.

Event: A competition within a discipline that results in a ranking and awards (e.g. medals).

Example of the relationship between **sport**, **discipline**, and **event** in the Olympic programme from the 2024 Paris Olympics:
- World Aquatics is the IF that governs the sport of aquatics
- Within the sport of aquatics are multiple disciplines – artistic swimming, diving, marathon swimming, swimming, and water polo.
- Within the discipline of diving are eight medal events:
  - Individual 3m springboard - men & women
  - Individual 10m platform - men & women
  - Synchronized 3m springboard - men & women
  - Synchronized 10m platform - men & women

References

[1] Olympics.com, https://olympics.com/en/paris-2024/medals

[2] Olympics.com Biography, Lang Ping, https://olympics.com/en/athletes/ping-lang

[3] USA Gymnastics Hall of Fame, https://usagym.org/halloffame/inductee/coaching-team-bela-martha-karolyi/

Disclaimer