The Olympic Games is a global sporting event where countries display their prowess. The prediction of the medal table has been one of the hot topics in the field of sports research and public attention, because the results of the medal table are affected by many factors, including the influence of famous coaches on the country's awards. Predicting the medal table is a challenging task, for which two models are developed: the LightGBM prediction model for Olympic medals and the multivariate linear regression model for the effect of famous coaches.

For Task 1, based on the LightGBM model. First, we processed the data provided by the topic, considered the influence of features such as NOC, Sport, Event, athletes\_number, Year, and Host on the prediction results, and calculated the correlation coefficient between each feature and the number of medals, and filtered out the features used for model training. The R^2 values of the model were 0.9885 for the number of gold medals, 0.9872 for the number of silver medals, 0.9804 for the number of bronze medals, and 0.9916 for the total number of medals. According to the adjustment of the program of the 2028 Olympic Games and the situation of the host country, the input data of the model were adjusted, and the awards of all countries were predicted. A confidence interval of 90% was set and the upper and lower limits of the prediction interval were calculated. Scores were assigned to the different medals, and the total country 2024 and 2028 medal scores were calculated to compare the advancing and retreating countries. Then, based on the 2028 predictions, we counted the countries that could win the first medal and estimated the probability of winning based on the number of predicted medals. Finally, we use AHP hierarchical analysis to calculate weights for the number of gold medals, silver medals, and bronze medals and calculate weighted scores based on the number of medals a country has won in a particular sport, to analyze which sports are most important for different countries.

For Task 2, based on a multiple linear regression model. First, we filter out the countries and corresponding Events that are most likely to have the “coaching effect”, and calculate the scores of these countries in the Olympic Games. Considering the lag of coaches' coaching, we added coach\_lag1-2 and coach\_lag3-5 as features for short-term (1-2 events) and long-term (3-5 events), and calculated the Pearson's correlation coefficient between each feature and the scores, and selected 10 features in total, and trained the short-term and long-term models using multiple linear regression model, and the R^2 value of the model was 0.828. Finally, we selected the United States, Japan, and the United Kingdom. We selected the United States, Japan, and the United Kingdom as the most likely countries to have the “famous coach effect”, and the corresponding Events, and calculated the scores of these countries. We chose three countries, the United States, Japan, and Great Britain, and compared the historical average scores on each sport with the scores of the last three Olympic Games to derive in which sports it is recommended to hire a famous coach.

For Task 3, we mined unique findings related to the Olympic medal standings from the two models, analyzed the relationship between medal distribution and program participation, medal standings and economic development level, as well as the training system and scientific research support behind the medals, and gave recommendations for the Olympic Committee to refer to. Finally, we conducted a sensitivity analysis on the parameters year and total\_event of the Olympic medal LightGBM prediction model. The high sensitivity of these parameters helps to capture the potential impact of subtle changes on the number of medals, leading to more accurate predictions.