

OLYMPIC ATHLETES DATA ANALYSIS

Tatevik Stepanyan, Maxime Fessy, Léo Le Terrien



1632

UNIVERSITY of TARTU
Institute of Computer Science

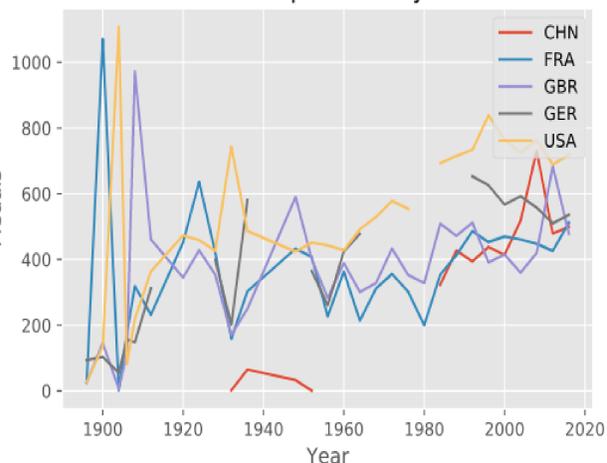
Introduction and Problematic

Country-teams participate in Olympic games on regular basis. Our objective is to demonstrate the trends of winning countries and the most popular sports. We are interested in sports that represent a source of further development and can attract the attention of sport-sector investors. We are also interested in the reasons of such opportunities.

Data Preparation

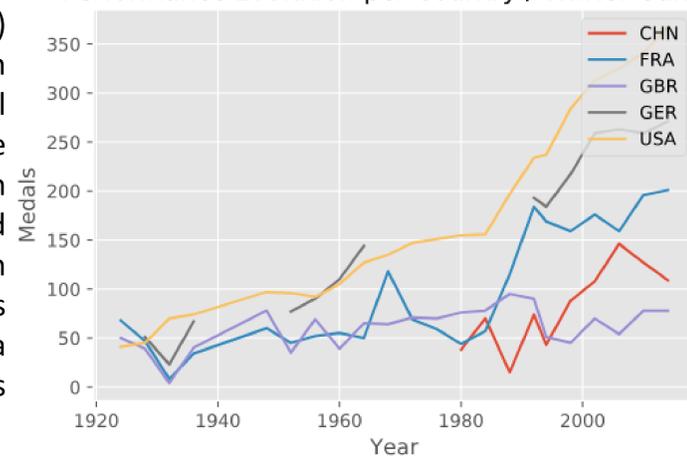
Olympic Athletes data is a Kaggle dataset. It comprises the historical data of athletes and medal results of Olympic Games from 1896 to 2016. The main analysis was conducted from the perspective of countries and sports in respect to medals.

Performance Evolution per Country / Summer Games

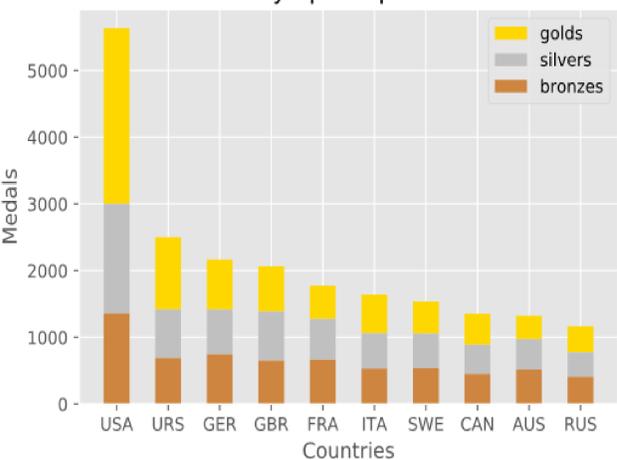


These graphs are about the evolution of number of athletes by country during the summer (left) and the winter (right) Olympics games. We can remark some gaps mainly due to political problems. At the beginning of the century, there was more athletes in summer Olympics than today, but we can also see that they are more and more athlete during the years in the main countries. USA had during almost all the games the more important number of athletes. China began really late to bring athletes to the games but today they caught up their late.

Performance Evolution per Country / Winter Games

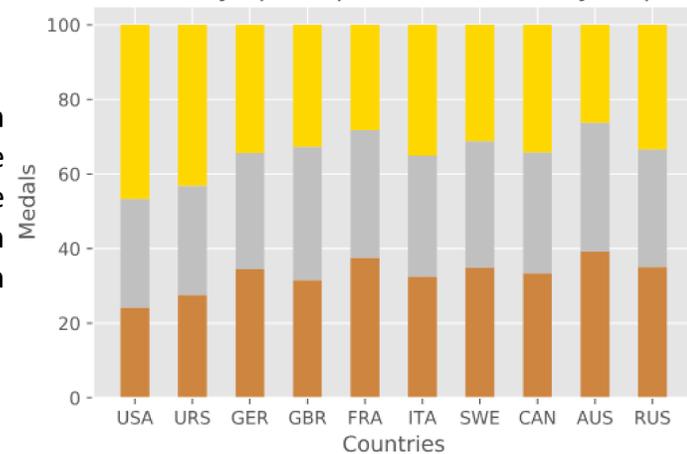


120-Years Olympic Top Teams' Medals



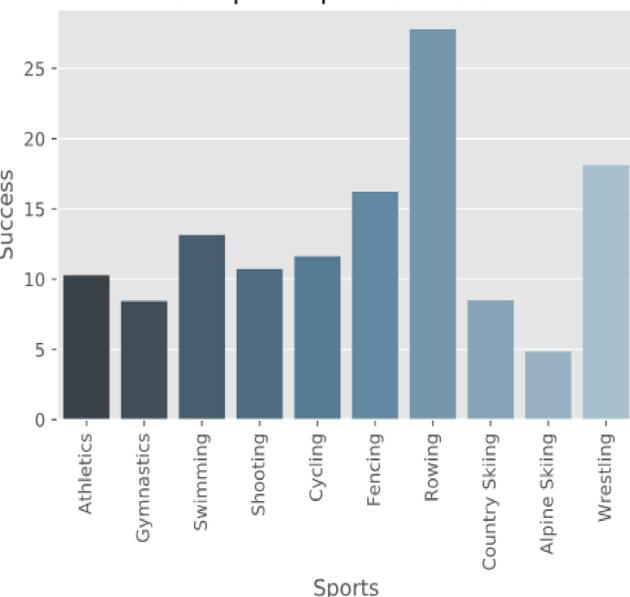
Overall, US remains the absolute leader in both the number of athletes taking part to the Olympic games and in the won medals. The composition of top 10 teams is stable in absolute numbers but also in the composition of the medals.

120-Years Olympic Top Teams' Medals by Proportion



The results show that the relative performance of top countries hide the outliers that are the main interest of our study. Success rate shows the proportion on medal winning athletes in the total number of participants for a given sport. Non surprisingly, the most popular sports record relatively low results explained mainly by strong competition and very limited places for medals. The results of the analysis revealed the top 15 sports that can be of high interest for further development given their specificity and lack of competition. On one hand, extremely high performance is conditioned by the fact that the sport can be represented by a single athlete (the case for aeronautics). On the other hand, even with relatively higher competition, the worst result of 20% are slightly behind the best result of 27% recorded for Rowing (the 7th most popular sport).

Most Popular Sports Success Rate



Can a "non-popular" sport be prioritized and provide benefits due to its specificity? According to our findings we suggest that those sports can be a source to develop and to ensure higher results with higher probability. What comes next is the question of the force of cultural and historical aspects of the sport and the personal preferences of athletes versus their will to succeed and the creativity of sport-sector investors to exploit new fields. The main question still remains: is it the lack of competition and the important niche that matters or is it the passion for the beloved sport that makes an athletes to win the gold.

Least Popular Sports Success Rate

