

# Olympic Games

Data Analysis

---

Clément GUERIN



# From data to viz

[ ] std\_scaler = StandardScaler()  
 X\_train = std\_scaler.fit\_transform(X\_train)  
 X\_test = std\_scaler.transform(X\_test)

Host versus non host medals

**EXÉCUTER** ENREGISTRER LA REQUÊTE

```

1 WITH HostMedals AS (
2     SELECT
3         UPPER(Athlete_country) AS Athlete_country,
4         COUNT(*) AS Host_Medals
5     FROM
6         `dulcet-aileron-411310.projet_jo.Medals_FINAL`
7     WHERE
8         UPPER(Host_country) = UPPER(Athlete_country)
9     GROUP BY
10        UPPER(Athlete_country)
11),
12 NonHostMedals AS (
13     SELECT
14         UPPER(Athlete_country) AS Athlete_country,
15         COUNT(*) AS NonHost_Medals
16     FROM
17         `dulcet-aileron-411310.projet_jo.Medals_FINAL`
18     WHERE
    
```



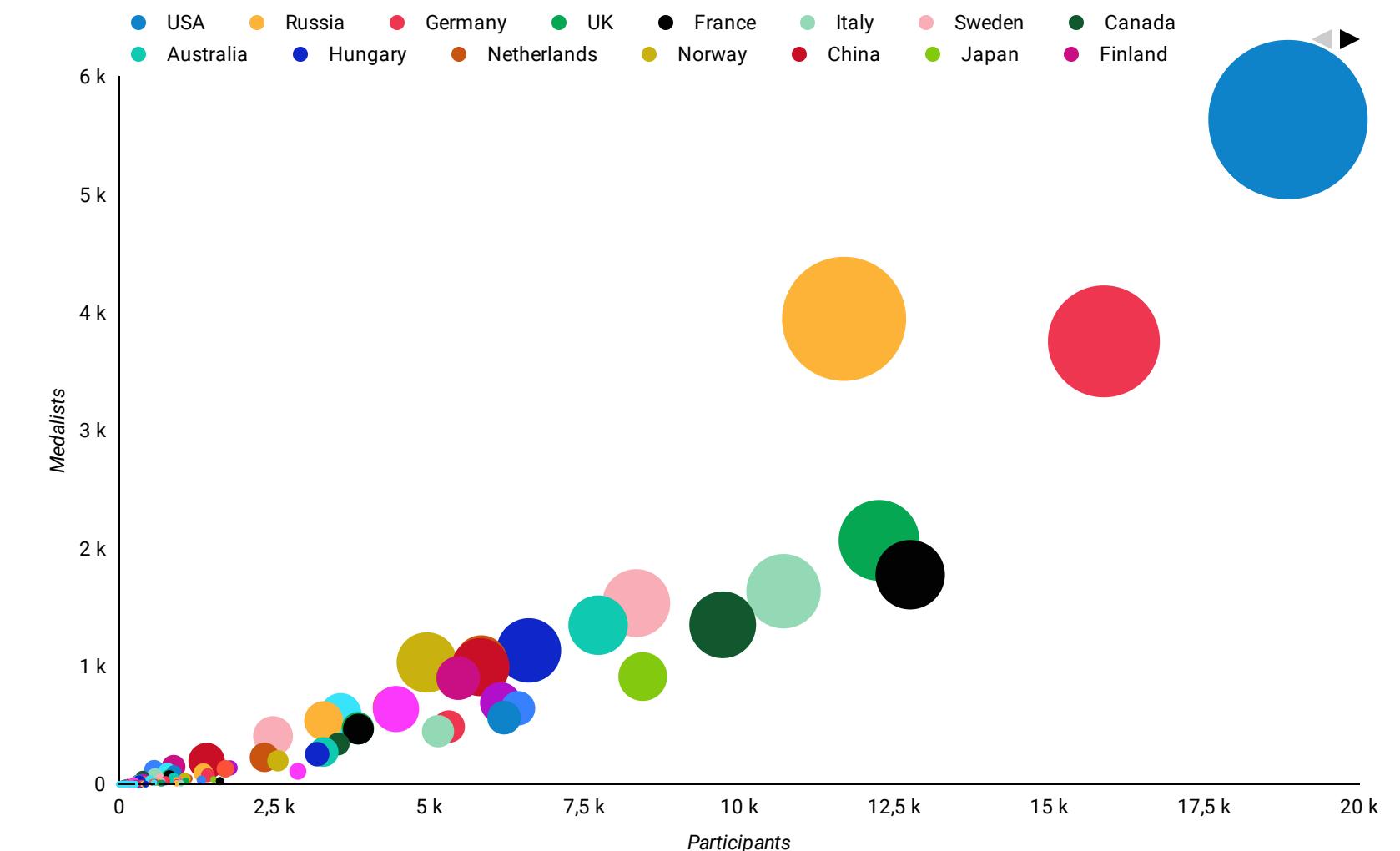
NonHost\_Medals

GRAPHIQUE JSON DÉTAILS DE L'EXI

pd.merge(ydf, zdf)  
 Rows that appear in both ydf and zdf  
 (Intersection).

pd.merge(ydf, zdf, how='outer')  
 Rows that appear in either or both ydf and zdf  
 (Union).

pd.merge(ydf, zdf, how='outer', indicator=True)  
 .query('\_merge == "left\_only"')  
 .drop(columns=['\_merge'])  
 Rows that appear in ydf but not zdf (Setdiff).





**Overview &  
evolution over  
time**

# Olympic Games from 1896 to 2022

---

Hosts  
**23**

Countries  
**238**

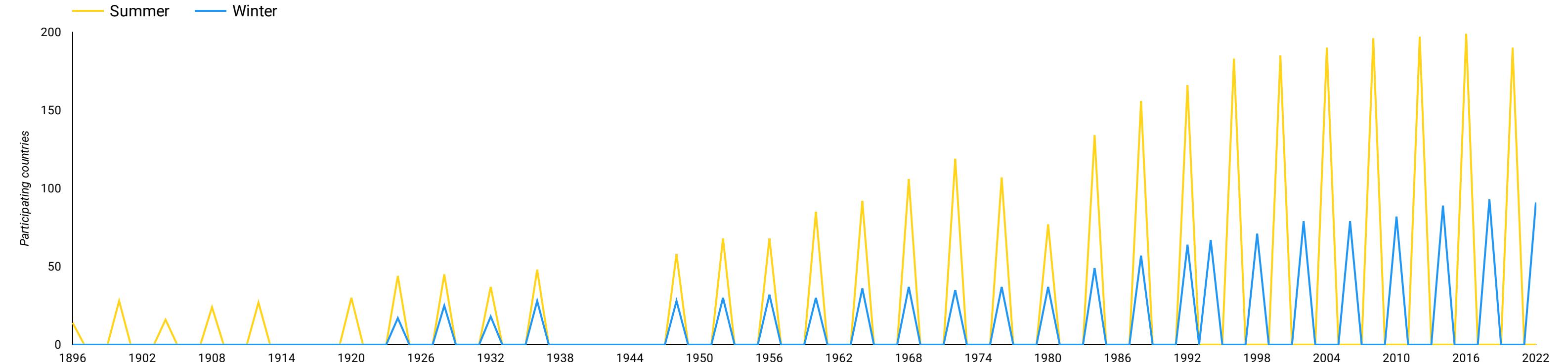
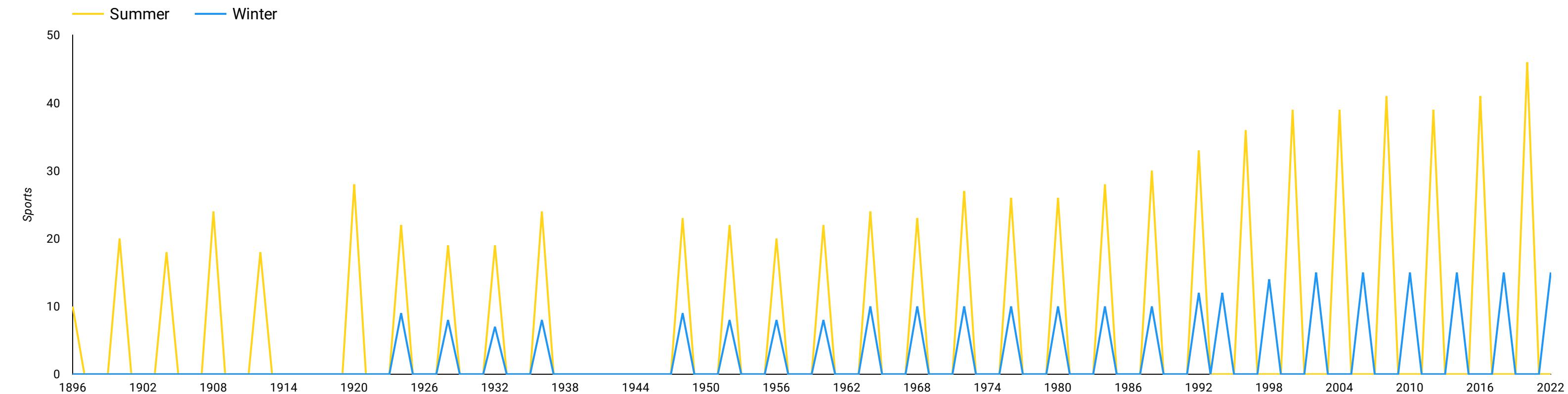


Editions  
**53**

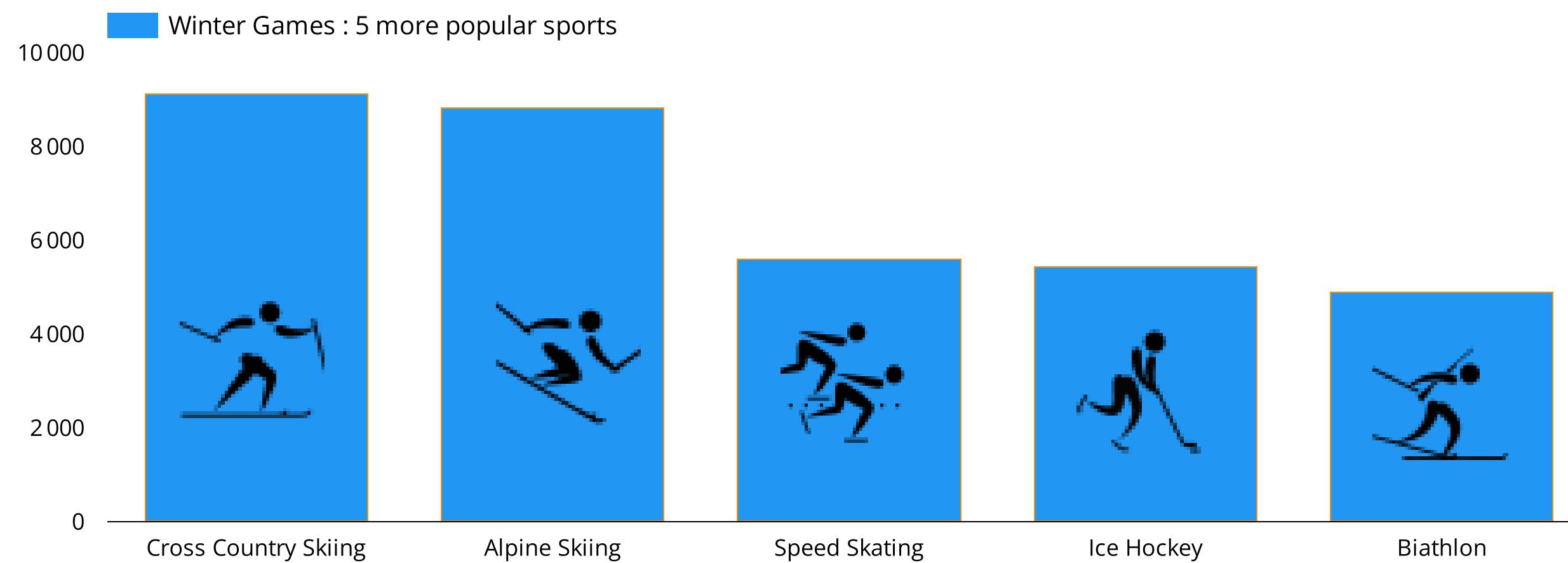
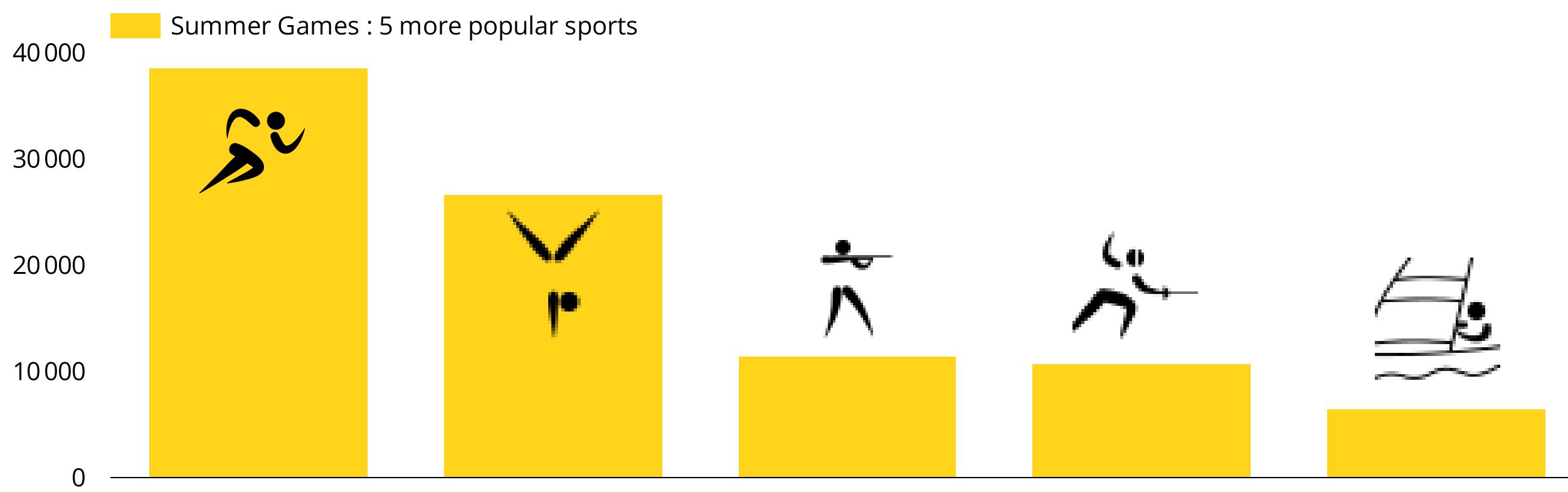
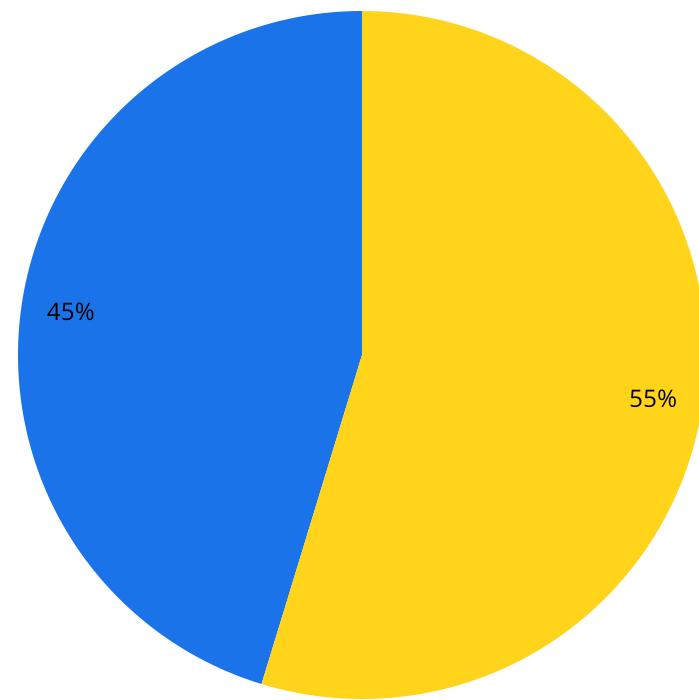
Sports  
**86**



# An increase in both sports & participants over the time



# Top 5 sports since 1896 : Summer vs Winter

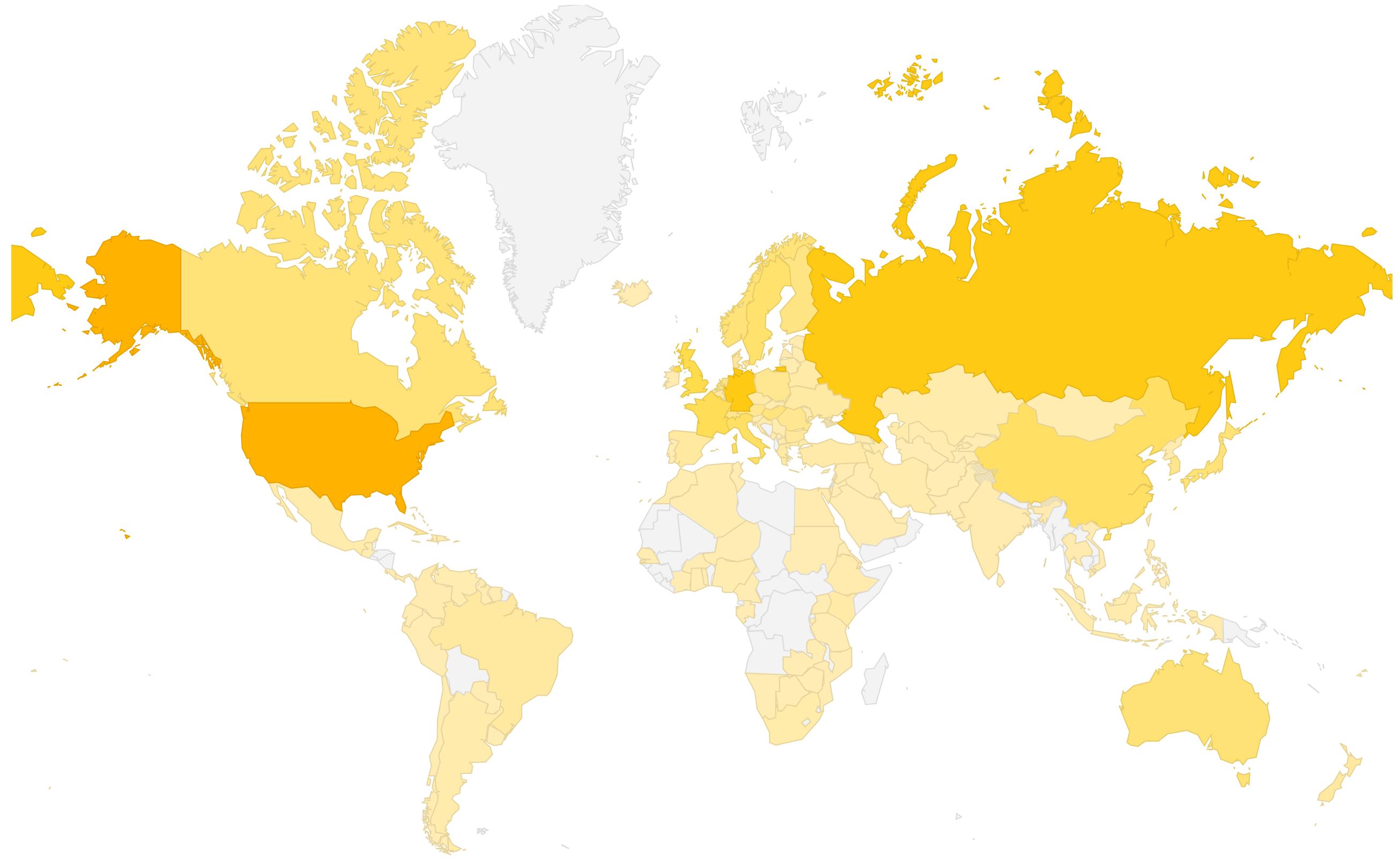




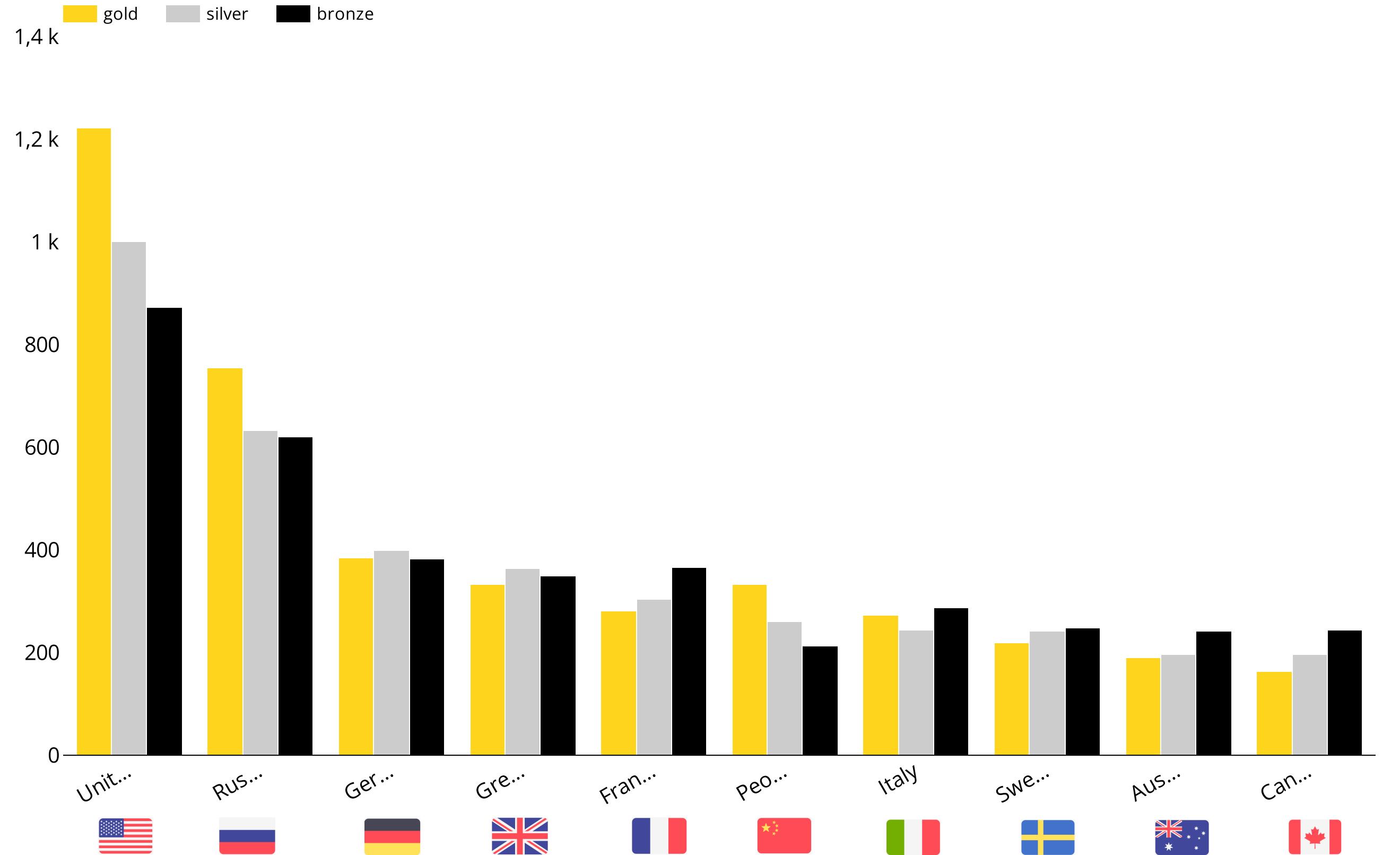
**International  
performance**

# Worldwide medals distribution

all Olympic Games (1896-2022)

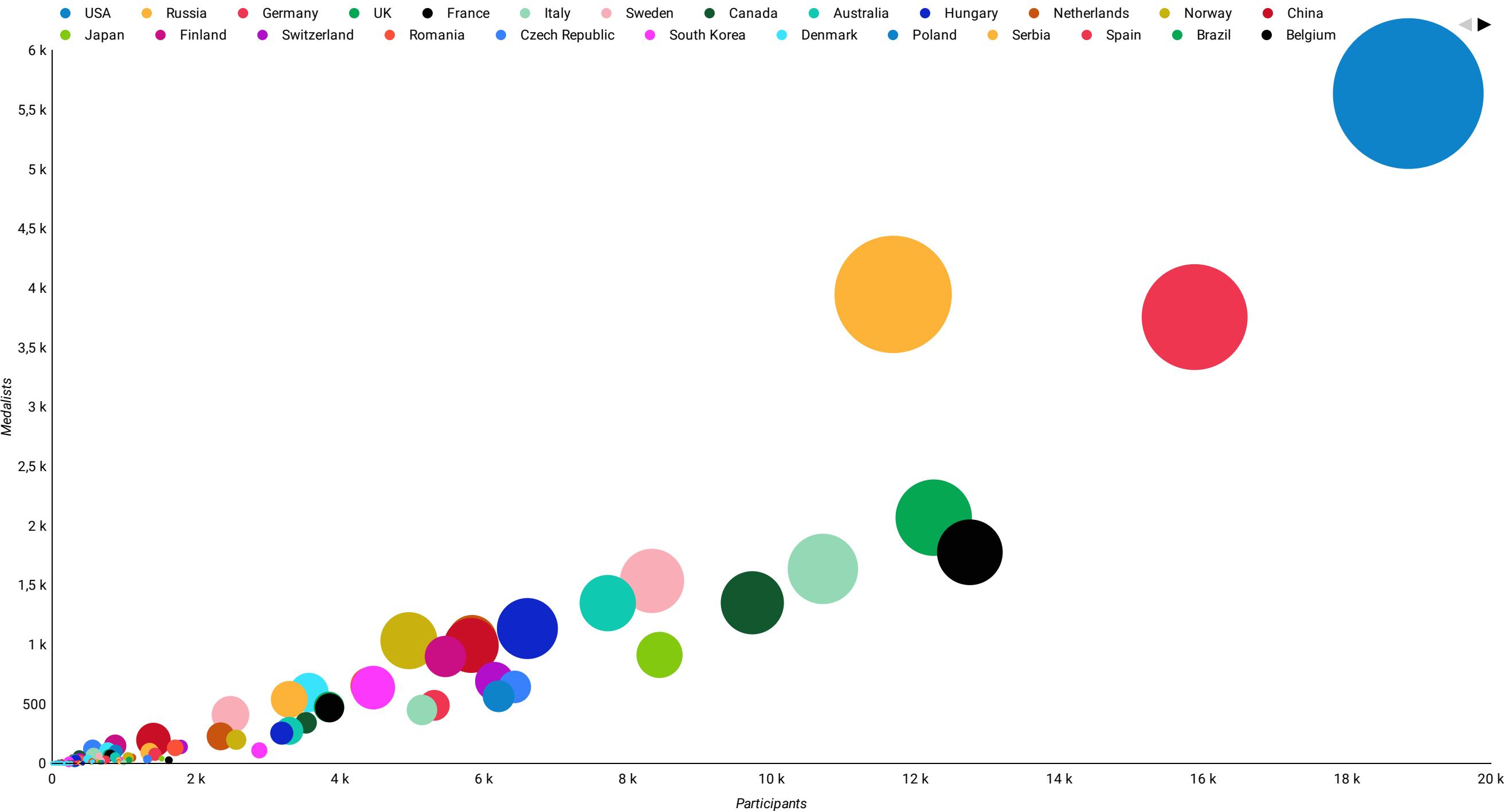


# Top 10 medalist countries



# Strong participation correlated to winning medals

3 overperforming countries who have a highest ratio of medals vs participations :  
Russia, Germany & USA.



# Top performing countries since 1896

—

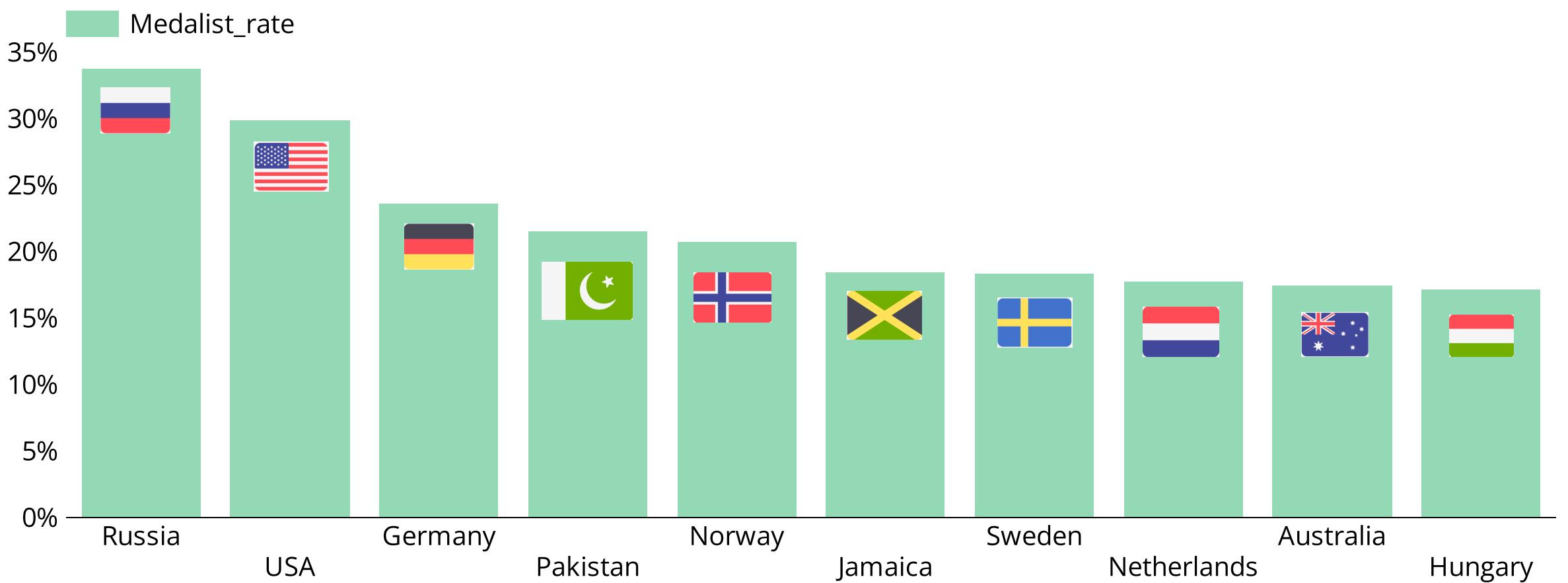
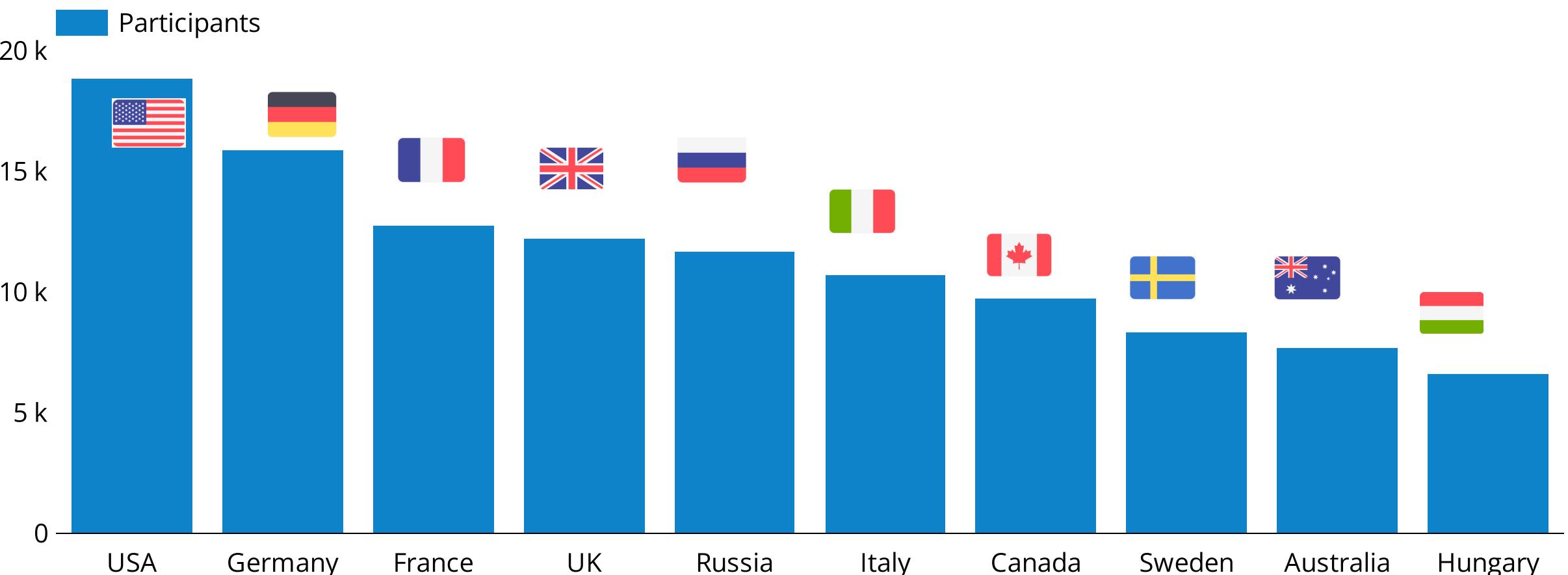
Example of Russia :

#5 in participation

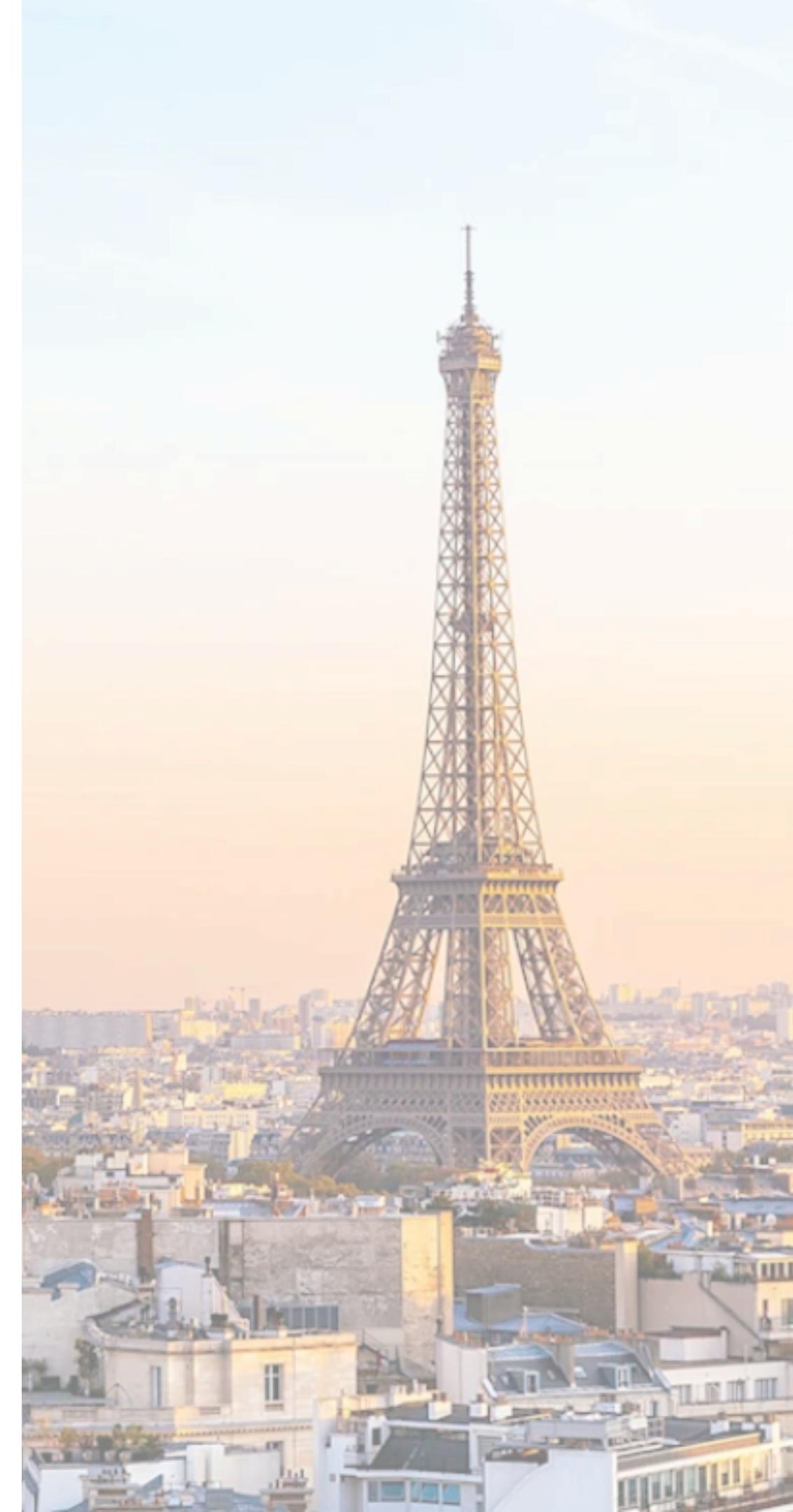
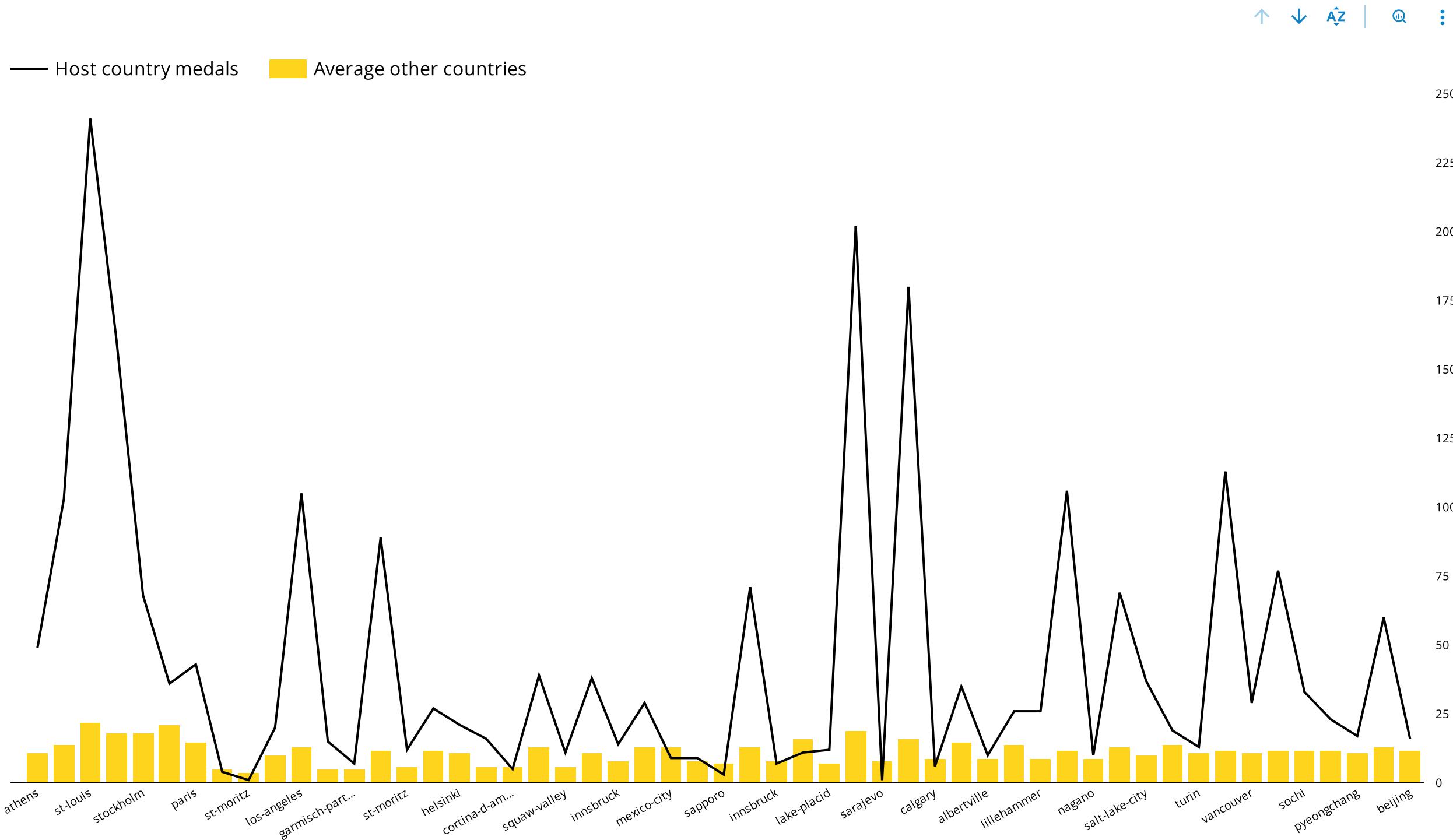
#2 in nb of medals

#1 in medal\_efficiency

France is not in the top 10  
anymore regarding the  
efficiency rate

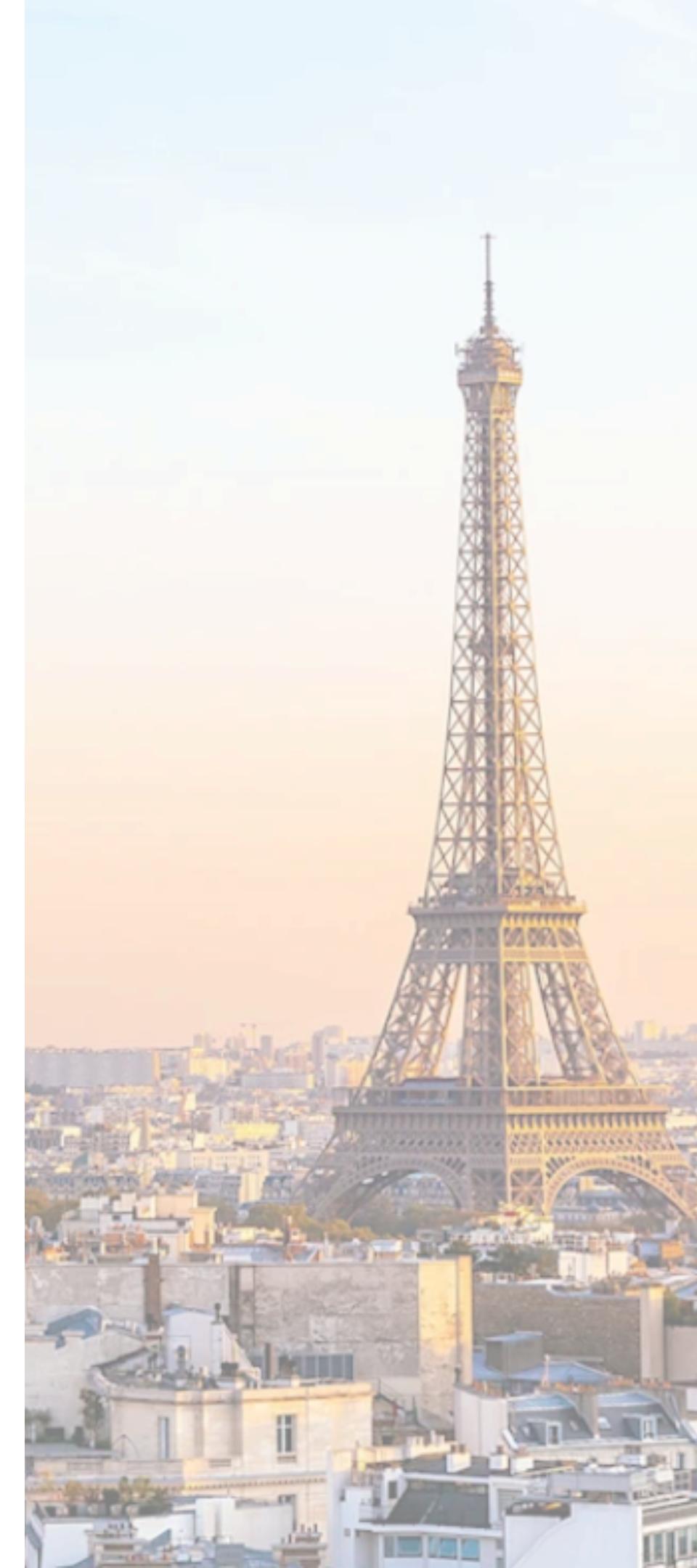
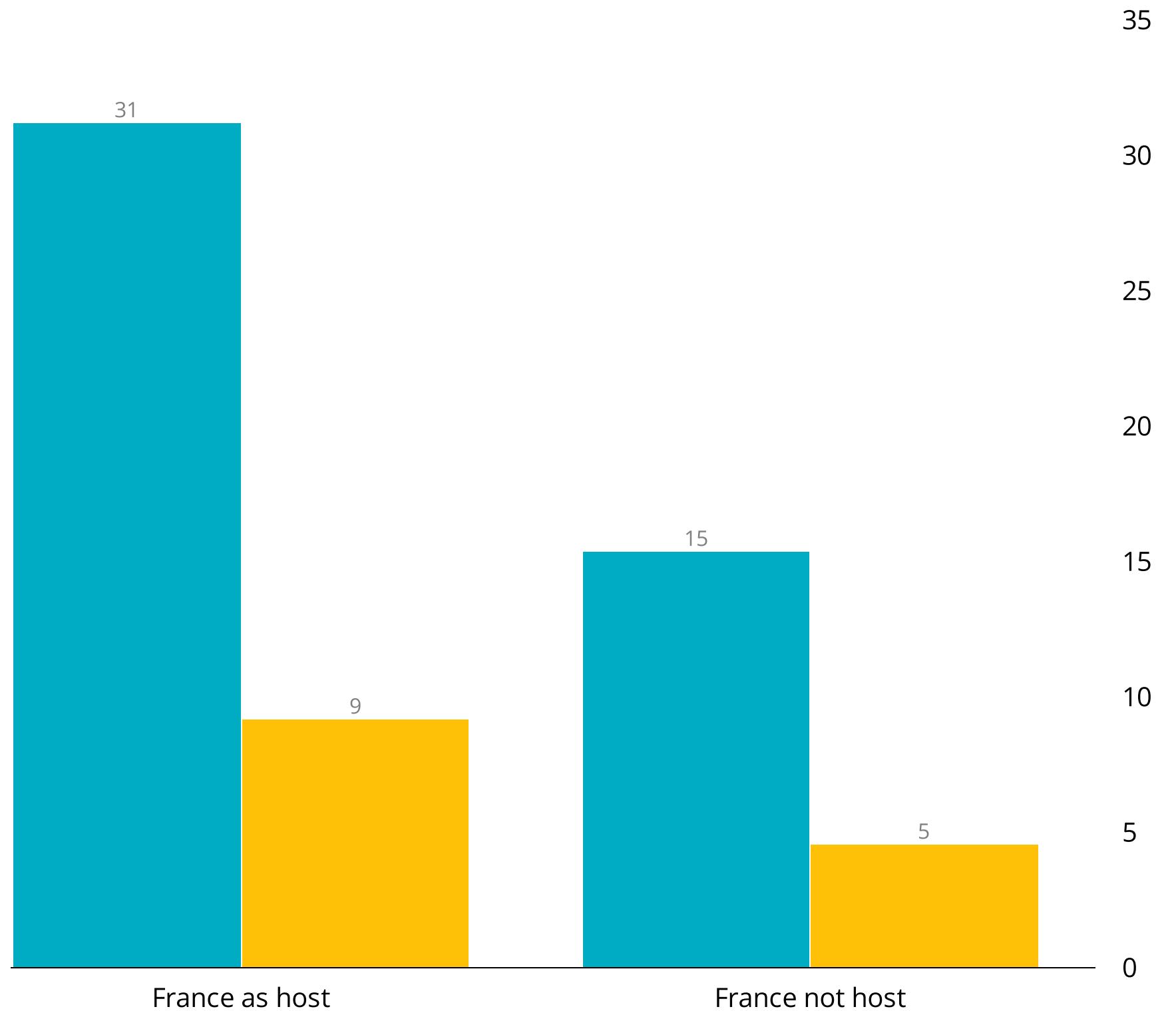


# Hosting : high impact on performance



# France : Hosting the Games bring 2x more medals

---



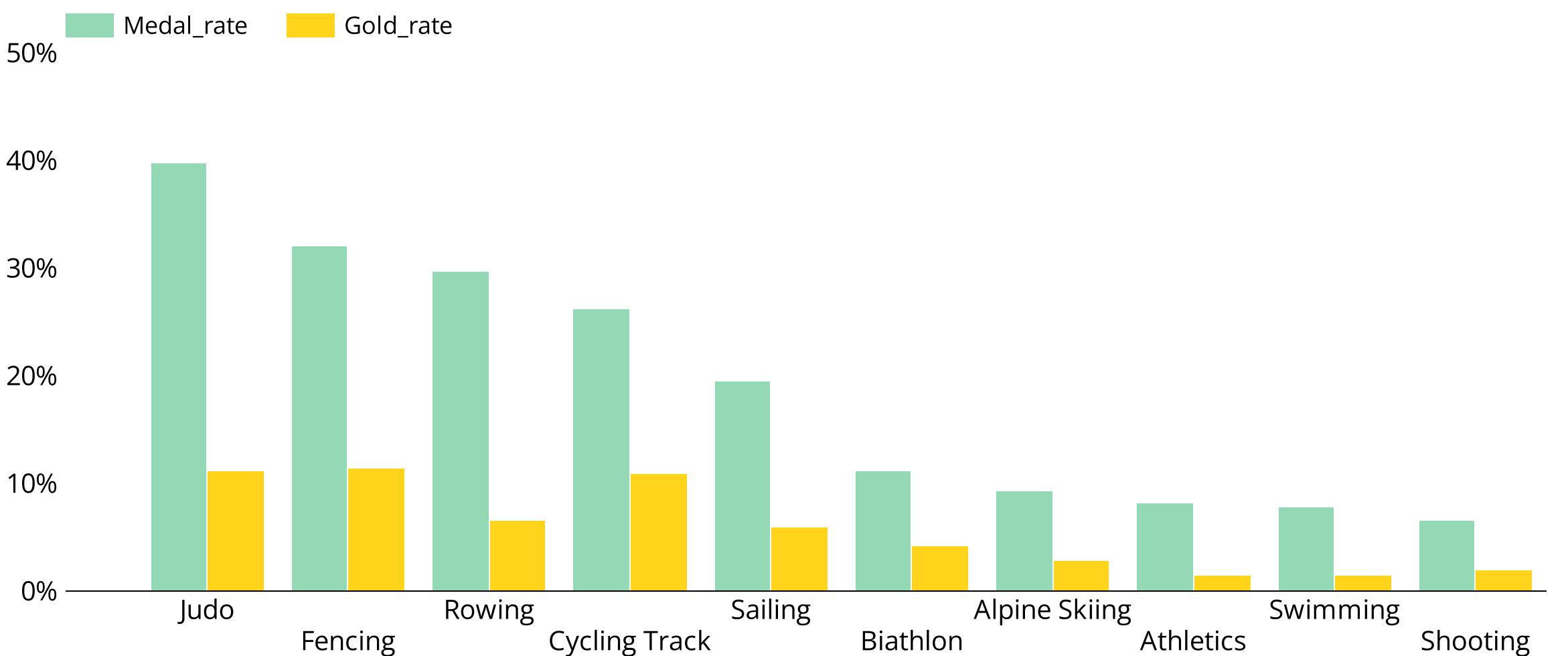
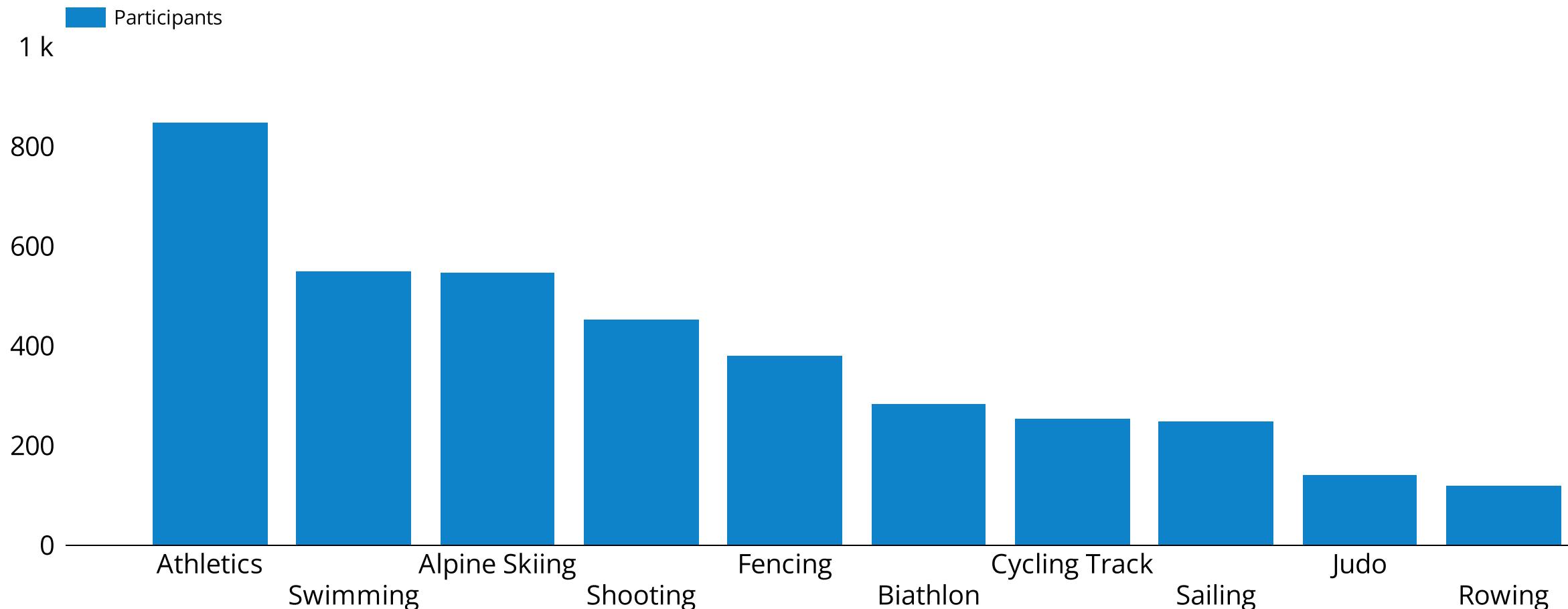


**Focus on France &  
prediction**

# France : Top 10 sports in terms of medals & titles

Athletics is the sport for which FR sends the most athletes though there is not a strong efficient ratio.

Judo & Fencing are the main medal & title drivers for France, followed by rowing, cycling track & sailing.



# **Paris 2024 : what about the medals to be won ?**

A correlation matrix shows which criteria seem more likely to influence the nb of medals that are to be won over.

The data that are more likely to influence the result are then taken into account in a prediction model.



# Paris 2024 : what about the medals to be won ?



(Paris 2024 : 52, Tokyo 2020 : 33)

# L'EQUIPE

Projection du nombre de médailles aux JO d'après l'ANS

Rang : 5e potentiel

Total



Our prediction

55  
medals



JO Paris 2024 23/01/2024 18:29

**JO de Paris 2024 : un top 5 de la France est-il possible comme le souhaite Macron ? Oui, selon notre projection**

À partir des résultats obtenus lors des championnats du monde disputés en 2023, la France pourrait remporter près de 60 médailles l'été prochain.

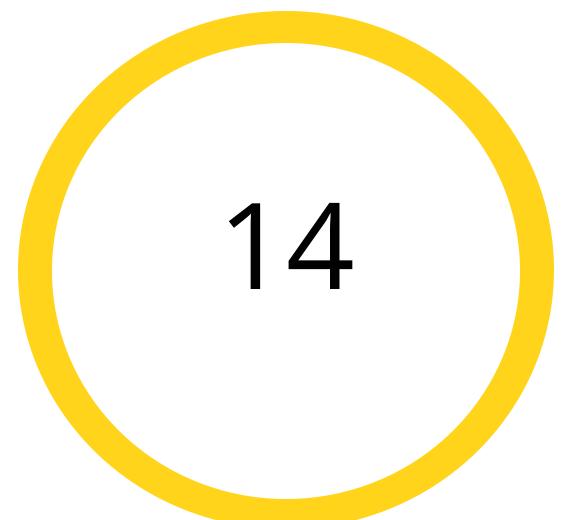
**Eurosport** REGARDER SCORES EN DIRECT FOOTBALL TENNIS CYCLISME TOUS

JO PARIS 2024 > VIDÉOS PARIS 2024 OLYMPIC CHANNEL MON PARIS OLYMPIQUE

**JEUX OLYMPIQUES PARIS 2024 : 54 MÉDAILLES POUR LA FRANCE, MARCHAND DEUX FOIS EN OR... LES PRÉDICTIONS DE GRACENOTE**

## Our predictive model : methodology

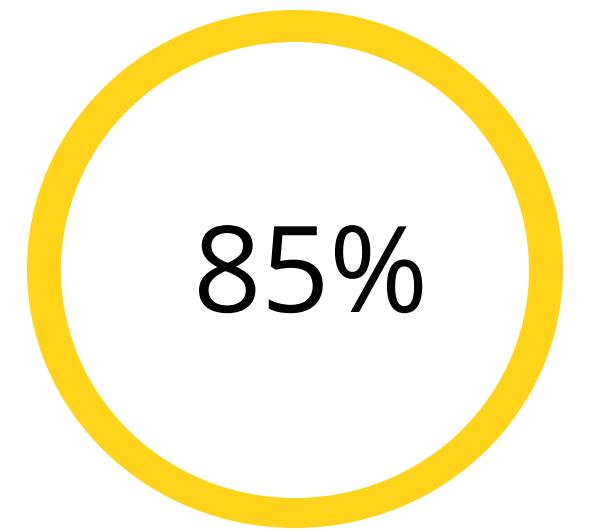
---



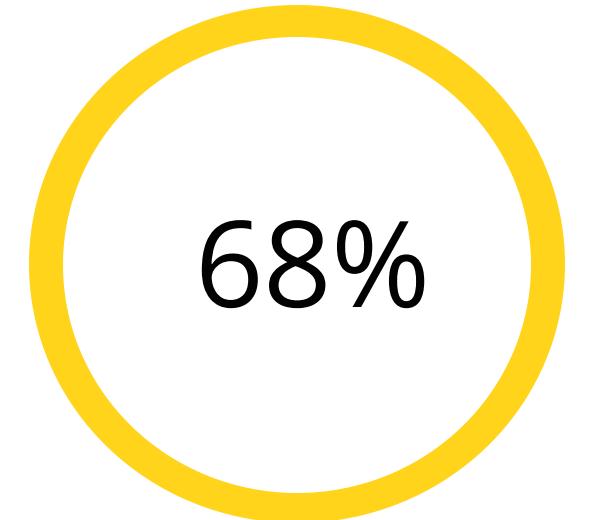
Mean  
Absolute Error



Predicted value



Summer / Winter ratio



Model score



# Hall of fame

Best world medalists



Michael Phelps



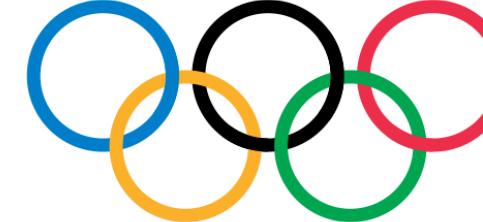
28



Larysa Latinyna



18



Best French medalists



8

Roger  
Ducret



8

Philippe  
Cattiau



Laura Flessel-Colovic



5



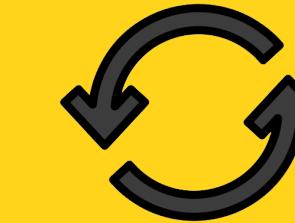
0.99 m

1.27 m - Rosario Briones,  
mexican gymnast (1968)  
2.26 m - Yao Ming, Chinese  
basketball player (2000)



189 kg

25 kg - Choi Myong-Hui,  
North Korean gymnast (1980)  
214 kg - Ricardo Blas Jr.,  
judoka from Guam (2008)



63 yr.

73 - Oscar Swahn, Swedish  
shooter  
10 - Dimitrios Loundras, Greek  
gymnast

An aerial photograph of the Seine River in Paris during the Olympic torch relay. The river is filled with small boats carrying people, and the banks are crowded with spectators. The Eiffel Tower is visible in the background against a sunset sky.

See you in July !